Glasgow
University Library

WITHDRAWN

L11 - a.23.

GEN. LANG.
C 205
MEI2

Book No 0286314

30114 002863143

Glasgow University Library

13. JAN. 1970

5. DEC. 74 C

17. FEB. 75 ◇

12 MAY 77 ◇

2 7 FEB 1986

GUL 68.18

COLLECTION LINGUISTIQUE

PUBLIÉE PAR

LA SOCIÉTÉ DE LINGUISTIQUE DE PARIS

A. MEILLET

I. LES DIALECTES INDO-EUROPÉENS

2e tirage ..... 9 fr.

II. MÉLANGES LINGUISTIQUES

offerts à M. F. DE SAUSSURE

1908 ..... 22 fr. 50

A. ERNOUT

III. LES ÉLÉMENTS DIALECTAUX DU VOCABULAIRE LATIN

1909 ..... *épuisé*

M. COHEN

IV. LE PARLER ARABE DES JUIFS D'ALGER

1912 ..... 50 fr.

M. GRAMMONT

V. LE VERS FRANÇAIS

SES MOYENS D'EXPRESSION, SON HARMONIE

2e édit. augm. 1913 ..... 30 fr.

DRZEWIECKI

VI. LE GENRE PERSONNEL DANS LA DÉCLINAISON POLONAISE

1918 ..... 14 fr. 40

SETALA

VII. LA LUTTE DES LANGUES EN FINLANDE

1920 ..... 5 fr.

A. MEILLET

VIII. LINGUISTIQUE HISTORIQUE ET LINGUISTIQUE GÉNÉRALE

2e édit. 1926.....

MAURICE CAHEN

IX. ÉTUDES SUR LE VOCABULAIRE RELIGIEUX DU VIEUX-SCANDINAVE

1921 ..... 36 fr.

MAURICE CAHEN

X. LE MOT « DIEU » EN VIEUX-SCANDINAVE

1921 ..... 14 fr. 40

JULES GILLIÉRON

XI. PATHOLOGIE ET THÉRAPEUTIQUE VERBALES. IV

1921 ..... 30 fr.

J. MAROUZEAU

XII. L'ORDRE DES MOTS DANS LA PHRASE LATINE

1922 ..... 36 fr.

JULES GILLIÉRON

XIII. THAUMATURGIE LINGUISTIQUE

1923 ..... 14 fr.40

A. CUNY

XIV. ÉTUDES PRÉGRAMMATICALES

SUR LE DOMAINE

DES LANGUES INDO-EUROPÉENNES ET CHAMITO-SÉMITIQUES

1924 ..... 60 fr.

A. MEILLET

XV. LE SLAVE COMMUN

1924 ..... 60 fr.

XVI. — LES LANGUES DU MONDE

1924 ..... 114 fr. — Relié ..... 134 fr.

XVII. — MÉLANGES LINGUISTIQUES

offerts à M. J. VENDRYES

1925 ..... 40 fr.

LOUIS RENOU

XVIII. LA VALEUR DU PARFAIT DANS LES HYMNES VÉDIQUES

1925 ..... 30 fr.

MARIE-LOUISE SJOESTEDT

XIX. L'ASPECT VERBAL ET LES FORMATIONS A AFFIXE NASAL EN CELTIQUE 40 fr.

ALBERT SÉCHEHAYE

XX. ESSAI SUR LA STRUCTURE LOGIQUE DE LA PHRASE

1926 ..... 36 fr.

P. CHANTRAINE

XXI. LE PARFAIT GREC. (*sous presse*)

GuL
0/16/57

LINGUISTIQUE HISTORIQUE
ET
LINGUISTIQUE GÉNÉRALE

COLLECTION LINGUISTIQUE

PUBLIÉE PAR

LA SOCIÉTÉ DE LINGUISTIQUE DE PARIS. -- VIII

LINGUISTIQUE HISTORIQUE ET LINGUISTIQUE GÉNÉRALE

PAR

A. MEILLET

membre de l'Institut

PROFESSEUR AU COLLÈGE DE FRANCE

DIRECTEUR D'ÉTUDES A L'ÉCOLE DES HAUTES ÉTUDES

DEUXIÈME ÉDITION

PARIS

LIBRAIRIE ANCIENNE HONORÉ CHAMPION, ÉDITEUR

5, QUAI MALAQUAIS, 5

—

1926

COLLECTION LINGUISTIQUE
PUBLIÉE PAR
LA SOCIÉTÉ DE LINGUISTIQUE DE PARIS. — VIII

LINGUISTIQUE HISTORIQUE

ET

LINGUISTIQUE GÉNÉRALE

A. MEILLET
membre de l'Institut
PROFESSEUR AU COLLÈGE DE FRANCE
DIRECTEUR D'ÉTUDES A L'ÉCOLE DES HAUTES ÉTUDES

DEUXIÈME ÉDITION

PARIS
LIBRAIRIE ANCIENNE HONORÉ CHAMPION, ÉDITEUR
5, QUAI MALAQUAIS, 5
1926

P. 297, note. *Lire* Quarterly, *au lieu de* Quaterly
P. 241, ligne 5 du bas, *lire* Père-Ciel, *au lieu* de Père-Cid

GLASGOW
UNIVERSITY
LIBRARY:

AVERTISSEMENT DE LA PREMIÈRE ÉDITION

Les chapitres qui constituent ce volume sont des articles qui ont paru, pour la plupart, depuis 1905, dans des périodiques divers. Quelques-uns ont été écrits pour le grand public, d'autres pour un public large mais curieux de science comme celui de la belle revue internationale *Scientia*, d'autres pour les philosophes ou les sociologues. Presque aucun n'a été destiné proprement à des savants dont la linguistique est la spécialité.

Ces articles n'ont subi que des retouches de détail. On y trouvera donc quelques redites, et des incohérences de forme.

Deux seulement, qui ont servi de thèmes à des conférences faites devant des publics très éclairés en Hollande, sont inédits.

Ecrits sans plan préconçu, ces exposés ont cependant une unité parce qu'ils se rattachent tous à quelques idées générales exposées dans le premier d'entre eux. C'est pour cela qu'ils sont rangés ici d'après les matières, et non par ordre chronologique.

Bien des gens croient pouvoir parler de langues sans avoir appris la linguistique. On souhaite que ce recueil leur fasse entrevoir l'extrême complication des faits et leur régularité, mais aussi la multiplicité des influences qui agissent sur les langues.

La grammaire a une méchante réputation près de bien des pédagogues. Comme la grammaire classique n'a guère suivi le progrès de la linguistique, elle se trouve en effet ne plus répondre à l'état actuel des idées. Les maîtres qui voudront bien lire ce recueil y apercevront peut-être le moyen de rendre parfois plus vivant et plus moderne l'enseignement de la langue, qui est une des tâches principales de l'école à tous ses degrés, et que, pourtant, on les prépare en général peu à donner.

Chaque siècle a la grammaire de sa philosophie. Le moyen âge a essayé de fonder la grammaire sur la logique, et, jusqu'au XVIII^e^ siècle, la grammaire générale n'a été qu'un prolongement de la logique. Le XIX^e^ siècle.

en étendant aux faits psychiques et sociaux de la méthode d'observation des faits qui est en usage dans les sciences physiques et naturelles depuis la Renaissance, a conduit à présenter la grammaire de chaque langue comme un ensemble de faits. Mais jusqu'ici ces faits ne se sont guère coordonnés. Les notes de cours de F. de Saussure, éditées sous le titre de *Cours de linguistique générale*, ont indiqué comment on y pourrait mettre un commencement d'ordre. Mais il reste à faire un grand travail pour ordonner les faits linguistiques au point de vue de la langue même.

L'objet propre de ce recueil est de montrer comment tout en obéissant à certaines règles générales que déterminent les conditions universelles de toutes langue, le changement linguistique est lié à des faits de civilisation et à l'état des sociétés qui emploient les langues considérées.

AVERTISSEMENT DE LA DEUXIÈME ÉDITION

Cette seconde édition ne comporte vis-à-vis de la première aucun changement. Le texte ancien a été reproduit mécaniquement. Il n'y a de nouveau que deux articles ajoutés à la fin.

L'ÉTAT ACTUEL

DES

ÉTUDES DE LINGUISTIQUE GÉNÉRALE

Leçon d'ouverture
du cours de Grammaire comparée au Collège de France
lue le mardi 13 février 1906.

En m'appelant à succéder à M. Bréal, au maître qui a introduit en France l'enseignement de la grammaire comparée, qui a le premier occupé cette chaire et pour qui elle semblait faite, le Collège de France et l'Académie des Inscriptions qui m'ont désigné, le gouvernement de la République qui m'a nommé m'ont fait un honneur dont je dois d'abord exprimer ma reconnaissance ; je ne m'étendrai pas davantage : la reconnaissance ne se traduit pas par des mots, et l'on mesurera la sincérité de mes sentiments au soin que je prendrai de la science dont les intérêts me sont confiés.

M. Bréal se retire au moment où il vient de publier ce livre sur la *Sémantique*, dont tout le monde a senti le charme et dont les progrès de nos études font chaque jour mieux apprécier la portée par les spécialistes ; il achève un nouvel ouvrage qui renouvellera une question capitale ; je n'ai pas à parler d'une carrière scientifique qui se poursuit avec éclat. Mais il est un trait de caractère que je tiens à indiquer, parce que j'en ai souvent éprouvé le bienfait, et dont la simple mention dira plus que beaucoup de louanges : M. Bréal a conseillé, soutenu et encouragé les jeunes gens sans leur demander de penser comme lui, et lorsque, après un enseignement long et glorieux, attristé seulement par la mort d'élèves éminents qu'il aimait, il a voulu abandonner sa chaire, il a souhaité d'y avoir pour successeur un disciple qui le continuerait en ne le répétant pas.

La jeune école linguistique française, dont une amitié fraternelle unit les membres, n'a présenté qu'un seul des siens à l'élection des corps savants ; celui qui occupe aujourd'hui cette chaire est donc aussi, en quelque manière, l'élu de ses collaborateurs, de ses anciens maîtres, de ses camarades, de ses élèves, et leur représentant : ce sera l'une des parties principales de sa tâche que de travailler à coordonner les efforts des linguistes français et que de contribuer à mettre en lumière leurs recherches. La leçon que vous allez entendre doit beaucoup de vues importantes aux travaux de l'ami qui a bien voulu faire associer son nom au mien sur la liste de présentation, M. Maurice Grammont, l'un de ces hommes rares qui apportent des idées nouvelles.

En entrant ici, j'y retrouve des maîtres à qui je dois beaucoup de ma formation intellectuelle, dont la chaude sympathie m'a soutenu dès le début de mes études et m'accompagne jusqu'à cette chaire : ils connaissent trop mes sentiments intimes pour me permettre de les remercier à cette place. Mais il en est un, mort prématurément, que j'ai la douleur de n'y plus rencontrer : James Darmesteter, à qui une brève carrière a suffi pour laisser une œuvre qui ne périra pas.

Je puis encore rappeler un autre nom : après avoir donné à notre pays dix ans d'un enseignement lumineux et avoir suscité autour de lui les vocations scientifiques, M. Ferdinand de Saussure est rentré dans sa patrie pour y occuper la chaire de grammaire comparée à la belle Université de Genève. Aucun de ceux qui ont eu le bonheur de les entendre n'oubliera jamais ces leçons familières de l'École des hautes études où l'élégance discrète de la forme dissimulait si bien la sûreté impeccable et l'étendue de l'information, et où la précision d'une méthode inflexiblement rigoureuse ne laissait qu'à peine entrevoir la génialité de l'intuition.

Mes dettes de reconnaissance sont immenses. Permettez-moi de ne les pas énumérer toutes et de commencer dès aujourd'hui l'enseignement de la grammaire comparée.

J'aborde immédiatement l'objet de ce cours, en vous exposant quelles considérations m'ont amené à rechercher, dans les pre-

mières leçons que je suis appelé à faire ici, les causes sociales des faits linguistiques.

I

Quand ils se proposent d'expliquer les changements qui surviennent dans les langues, les linguistes recourent d'ordinaire à un très petit nombre de notions fondamentales distinctes.

Ils constatent que, à certains moments, en certains lieux, la prononciation subit telle ou telle modification, et ils formulent en une *loi phonétique* cette modification qui atteint la prononciation en tant que telle, et indépendamment de toute considération de sens ou de rôle grammatical. Ainsi, à la fin du mot, dans la France du Nord, un *a* de syllabe finale devient *e* muet entre le VI^e^ siècle et le X^e^ après Jésus-Christ ; un impératif *canta* devient *chante*, exactement comme un féminin *lenta* devient *lente*. — D'autres changements ont lieu en fonction du sens, du role grammatical ou syntaxique, et, bien qu'ils puissent à l'occasion avoir des résultats identiques aux précédents, ces changements dits analogiques en diffèrent par leur nature d'une manière essentielle. Ainsi *cantas*, *cantat* sont devenus (*tu*) *chantes*, (*il*) *chante*(*t*) par de simples changements de la prononciation ; et les changements de la prononciation ont simultanément transformé *canto* en (*je*) *chant* (sans *e* muet final), comme ils ont transformé *homo* en *on* ; le français du XI^e^ siècle conjuguait donc *chant*, *chantes*, *chantet*, mais certains verbes du même type, tels que *trembler* ou *entrer*, avaient *e* muet à la I^re^ personne, soit *tremble*, *entre* ; sur ce modèle et sous l'influence concomitante de *chantes*, *chantet*, l'*e* muet s'est introduit à la I^re^ personne, et l'on a dit *chante* ; on le voit, l'*e* muet de *tu chantes* résulte d'un changement de la prononciation, celui de *je chante* d'un changement de la flexion grammaticale ; ce sont deux procès distincts, et qui n'ont qu'un trait commun, celui de s'être développés spontanément et sans aucune influence extérieure appréciable.

Mais, et ceci est un troisième type distinct des deux précé-

dents, la langue d'une localité donnée est toujours plus ou moins accessible à l'influence des populations avec lesquelles ses membres sont en rapport : on emprunte des mots, des tours syntaxiques, des formes grammaticales, des manières de prononcer à d'autres langues, à d'autres parlers ou même à des textes écrits. Le résultat peut être exactement pareil à celui des changements spontanés ; ainsi la prononciation *wa* s'est substituée à *wè* à Paris ; cette prononciation *wa* est reproduite actuellement dans tous ceux des parlers français qui subissent l'influence parisienne ; *loi* s'y prononce donc *lwa* comme à Paris, et non plus *lwè* ; mais à Paris le changement était spontané, tandis que, dans ces parlers, *wa* est emprunté. Le résultat est identique, mais le procès diffère absolument, et c'est le procès qu'il importe avant tout d'analyser.

Les « lois phonétiques », l'analogie, l'emprunt, tels sont les trois principes d'explication qu'a reconnus la linguistique au cours du XIX^e^ siècle ; appliqués à des langues très diverses, de tous temps et de tous pays, ils se sont partout et toujours trouvés vérifiés par l'expérience ; et l'on s'est rendu mieux compte de l'histoire des langues à mesure qu'on les a employés avec plus de rigueur et de précision, et qu'on a suivi de plus près, analysé avec plus d'exactitude les changements de prononciation, les innovations analogiques et les emprunts de toutes sortes. La linguistique en a été renouvelée tout entière.

Pour immense que soit le parti qu'on en a tiré, on en peut attendre plus encore qu'ils n'ont déjà fourni à l'explication des langues.

Tout d'abord, il reste des familles entières de langues auxquelles on n'a presque pas commencé de les appliquer, fût-ce d'une manière élémentaire, et même dans les groupes où l'application en a été poussée le plus loin, il reste une infinité de parlers, de dialectes, de langues même où presque tout est encore à faire.

Dans les domaines qui ont été étudiés le plus attentivement, il n'y a sans doute pas une question dont on ne puisse renouveler l'étude en apportant à l'application des principes une pré-

cision nouvelle, en utilisant les découvertes des philologues qui se consacrent à chaque langue, en refaisant le travail philologique et en reprenant l'étude de tout l'ensemble des faits qui, de près ou de loin, se rapportent au sujet ; en linguistique, comme en toute science, la solution de beaucoup de problèmes tient à un degré de précision de plus dans la détermination des faits ; les solutions qui semblent acquises deviennent incertaines quand on serre de près les dépouillements sur lesquels elles reposent ; et chaque précison obtenue vient poser des problèmes nouveaux.

Aucune philologie ne se suffit à elle-même. Et, si l'indianiste est obligé de demander à la Grèce et à la Chine les dates que les textes sanskrits lui refusent, le linguiste qui étudie une langue est souvent aussi obligé de demander des témoignages à la philologie des langues étrangères : ce sont les mots arméniens empruntés à l'iranien qui ont permis d'écrire la phonétique historique du persan.

Et ce n'est pas de la philologie seule que le linguiste attend des précisions plus grandes : le temps n'est plus où la linguistique était un département de la philologie, et où la grammaire comparée pouvait recevoir parfois un nom que les linguistes n'ont d'ailleurs jamais adopté, celui de philologie comparée. L'observation des faits actuels est encore plus capable d'expliquer le passé que l'étude du passé d'expliquer le présent, et les langues modernes, tant dans leurs formes les plus populaires que dans leurs formes écrites et littéraires, ont attiré l'attention des savants, qui se dirigeait autrefois d'une manière trop exclusive sur les langues qu'on ne peut plus observer ; on s'est mis à décrire avec une exactitude singulière tous les détails des idiomes modernes.

On ne se contente même plus de l'observation directe, et l'on s'est ingénié à inventer des appareils qui permettent d'enregistrer les sons émis et d'inscrire chacun des mouvements articulatoires ; les premiers essais de ce genre ont été faits ici, dans le laboratoire de physiologie ; reprises depuis en des proportions plus étendues, ces recherches ont amené la création d'un labo-

ratoire de phonétique à côté de la chaire de grammaire comparée. La mesure s'introduit ainsi dans la phonétique, et c'est le commencement d'une petite révolution.

Moins encore peut-être que de l'observation phonétique directe on peut se contenter des remarques vagues et fuyantes de la psychologie vulgaire, avec lesquelles les linguistes ont trop longtemps opéré. En cherchant à tirer parti des faits linguistiques, les psychologues ont été conduits à les éclairer à l'aide des données de la psychologie moderne ; or, il n'y a pas de fait linguistique qui ne repose sur quelque activité psychique, et dans l'étude duquel on ne puisse profiter des découvertes de la psychologie. Quelques linguistes sont même allés jusqu'à vouloir trouver dans la psychologie l'explication de tous les faits linguistiques ; c'est une grave erreur, mais qui procède d'un point de départ juste.

En même temps l'étude du vocabulaire se renouvelle, et renouveler l'étude du vocabulaire, c'est renouveler toute la phonétique historique, qui repose sur l'examen étymologique des mots, et par là toute la linguistique historique. D'une part on a compris que l'étude des mots ne peut se séparer de l'étude des choses désignées par ces mots ; de l'autre, les atlas linguistiques qui se préparent de divers côtés et dont la publication a même commencé — la France a pris de ce côté une remarquable avance — fournissent à l'étude du vocabulaire des outils de recherche dont les premiers résultats acquis font entrevoir la décisive importance. Quand on constate l'existence d'un mot en latin et de son représentant phonétiquement correct dans un parler français moderne, on est au premier abord tenté de croire que ce mot s'est simplement transmis de génération en génération ; la géographie linguistique, combinée avec l'examen des choses et l'histoire des choses, a montré que cette vue simple était une vue inexacte ; elle a révélé des séries d'emprunts dans des cas où l'on supposait, assez naïvement, la persistance d'un même vocable durant des suites illimitées de siècles. Il apparaît de plus en plus qu'on s'est exagéré le rôle du changement spontané ; on a attribué au changement spontané, phonétique ou morpholo-

gique, tout ce que l'on a pu expliquer par là, et l'on se plaisait à ne voir dans l'emprunt qu'un fait accessoire ; en réalité, l'emprunt est un fait normal, et dont l'importance dans le développement linguistique éclate chaque jour davantage.

Ainsi, de toutes parts, à l'application schématique des trois principes de la « loi phonétique », de l'action analogique et de l'emprunt on voit se substituer l'observation toujours plus précise de réalités toujours plus complexes et plus variées. Et, dans la suite de ce cours, il y aura lieu de mettre en évidence cet enrichissement. Mais si près de la réalité que permettent d'approcher les progrès de la philologie, de la physiologie, de la psychologie, de la géographie linguistique, de l'étude des choses elles-mêmes, et si soigneusement que les linguistes tiennent compte de la complication souvent inextricable des faits, le défaut essentiel de toute méthode historique demeure : malgré toutes les précisions, malgré tous les enrichissements, les principes posés n'expliquent jamais que des faits particuliers, et ne fournissent que des conclusions particulières ; on aboutit à une poussière d'explications, dont chacune est juste peut-être, mais qui ne constituent pas un système, et qui ne sont pas susceptibles d'en constituer jamais un. La constitution de l'histoire des langues a été un moment essentiel dans le développement de la linguistique ; mais l'histoire ne saurait être pour la linguistique qu'un moyen, non une fin.

II

Le développement linguistique obéit à des lois générales. L'histoire même des langues suffit à le montrer par les régularités qu'on y observe.

En effet, quand on examine les changements qu'a subis la langue indo-européenne commune sur les divers sols d'Asie et d'Europe sur lesquels elle s'est implantée, on fait une double constatation. Pour le détail matériel des changements, chaque dialecte a suivi ses voies propres, si bien qu'aujourd'hui les diverses langues indo-européennes ont des systèmes phoné-

tiques, des grammaires, des vocabulaires entièrement distincts, et que les traces de leur ancienne unité sont ou tout à fait indiscernables ou sensibles seulement à un spécialiste exercé. Mais, en même temps, ces changements, tous différents les uns des autres dans leur matérialité, sont exactement semblables dans leur direction générale.

A l'égard de la prononciation, les articulations qui ont été altérées sont presque partout les mêmes. Ainsi, et ceci est un principe général, la fin des mots a souffert plus que les initiales. Même une langue, dont l'état de conservation, sur ce point comme sur tant d'autres, émerveille le linguiste, le lituanien, impose aux voyelles finales de ses mots des abrègements et des changements de timbre dont l'intérieur du mot n'a pas l'équivalent. Les autres langues indo-européennes, moins conservatrices, ont toutes perdu plus ou moins complètement la fin du mot indo-européen ; les rares mots indo-européens qui se retrouvent en français n'en ont plus aujourd'hui la moindre trace, sinon dans l'écriture, du moins dans la prononciation ; un indo-européen **esti* est devenu *è* dans *il est*, **donom* est devenu *don*, l'accusatif féminin **oinām* est devenu *une*, dont l'*e* muet final, dernière trace de l'*a* latin, ne se prononce plus, et ainsi dans tous les cas. C'est donc une tendance générale des langues indo-européennes que la tendance à articuler les fins de mot d'une manière particulièrement débile.

Le développement morphologique des langues indo-européennes présente des tendances générales non moins nettes. Ainsi la flexion indo-européenne commune était très compliquée et comprenait un grand nombre de formes diverses ; les rapports que les mots soutiennent entre eux dans la phrase étaient indiqués par des formes flexionnelles variées, et par suite la phrase indo-européenne se présentait comme un agrégat très lâche d'éléments autonomes, rangés dans un ordre libre ; cet ordre dépendait seulement de l'importance attribuée à telle ou telle notion par le sujet parlant. A mesure que les langues de la famille ont évolué, toutes ont, plus ou moins tôt, plus ou moins vite et plus ou moins complètement, réduit l'importance de la

flexion et resserré plus étroitement les éléments de la phrase. Là où l'indo-européen avait trois modes, l'indicatif, le subjonctif et l'optatif, que distinguent encore toutes les langues attestées sous la forme la plus archaïque, le grec ancien, l'iranien ancien, le sanskrit védique, on n'en trouve bientôt plus que deux, comme en latin, en irlandais, en germanique, en arménien, ou même plus qu'un seul, comme en slave ; de même qu'en latin, il n'y a que deux modes en grec moderne par contraste avec le grec ancien, en sanskrit classique et en prâkrit par contraste avec le védique. De même, les huit cas de la déclinaison indo-européene n'apparaissent plus que dans les formes anciennes de l'indo-iranien ; des autres langues, celles qui ont le plus conservé ont perdu un cas, comme l'arménien, le polonais, le lituanien, le latin ancien, ou deux, comme le russe ; et, sauf une amorce d'illatif en lituanien oriental, on ne voit pas qu'aucun cas nouveau ait été ajouté à ceux que distinguait l'indo-européen commun. Les relations des mots entre eux et les nuances de sens exprimées par les cas ont été rendues par d'autres procédés : par l'ordre des mots qui tend à devenir fixe, de libre qu'il était, et par des mots spéciaux : prépositions, conjonctions, articles. Ce développement est nécessaire : dans une langue comme l'arménien moderne, où la flexion nominale a des formes distinctes pour un nombre de cas à peine moindre que celui de l'indo-européen, les désinences qui marquent chaque cas sont si fixes et si constantes, identiques d'ailleurs pour le singulier et pour le pluriel, qu'elles sont de tous points comparables aux prépositions françaises ; une observation pure et simple des faits, qui ne tiendrait pas compte de l'histoire — et c'est ainsi qu'on devrait toujours décrire les langues — aboutirait à mettre sur un même plan les unes et les autres ; en même temps, le fait que le nominatif et l'accusatif ont presque toujours une seule et même forme en arménien moderne a pour conséquence immédiate un ordre fixe des membres nominaux de la phrase verbale, comme en français.

Les développements phonétiques et morphologiques des langues indo-européennes, divers dans leur détail matériel, ont

donc obéi à des tendances exactement semblables, et présentent un saisissant parallélisme.

De là résulte, pour le dire en passant, la nécessité, que trop longtemps on n'a pas aperçue, de suivre dans toute son étendue, depuis l'indo-européen jusqu'à l'époque moderne, la courbe du développement de chacune des langues de la famille. Les comparatistes ont cru longtemps, quelques-uns croient peut-être encore, qu'on peut se contenter d'expliquer les formes les plus anciennes de chaque langue, en les rapprochant du type indo-européen ; c'est un procédé commode, et qui permet d'ignorer beaucoup de choses, ce qui est utile, mais il est artificiel. Qui veut vraiment expliquer n'a pas plus le droit d'isoler les périodes modernes des périodes anciennes que l'on n'a le droit d'expliquer l'état actuel par lui-même, en négligeant le passé. Le latin n'est qu'un moment de la grande transformation qui partant de l'indo-européen a abouti aux parlers romans actuels et qui de ceux-ci aboutira à quelque état nouveau. Ainsi les altérations qui de l'indo-européen **esti* « il est » ont fait le français *est*, l'italien *è* avaient commencé dès avant le latin historique, où l'on a déjà *est*, et non plus **esti* ; le français *une*, qui représente un indo-européen **oinām*, était déjà préparé en latin par l'abrègement de l'*ą* de *unam* et par l'affaiblissement très marqué de l'articulation de la nasale finale. De même l'échelonnement des altérations par lesquelles les huit cas de l'indo-européen ont disparu en français est remarquable : dès avant la période historique du latin, dès l'italique commun sans doute, l'ablatif et l'instrumental sont déjà fondus en une forme unique ; l'ancien latin a encore un vocatif distinct du nominatif, mais seulement au singulier et seulement dans la seconde déclinaison ; il a encore un locatif *Karthaginī* distinct de l'ablatif *Karthagine*, mais seulement au singulier, et dans certaines conditions strictement définies ; ce locatif tend à disparaître à l'époque classique ; les autres cas se maintiennent tant que dure le latin littéraire ; quand les dialectes romans de la Gaule commencent à être écrits, on n'y reconnaît plus que deux cas ; c'est au XIV^e^ siècle seulement, à la date où toute déclinaison disparaît en français, qu'on

peut dire que la ruine de la déclinaison indo-européenne, commencée bien avant l'époque historique, est achevée dans les dialectes italiques. Par une conséquence nécessaire, l'ordre des mots est devenu de plus en plus fixe, développement qui ne s'est terminé qu'à une époque moderne. — Il y a donc une continuité dans l'évolution linguistique, et cette continuité révèle la constance des causes qui déterminent les modalités du changement.

Les changements linguistiques ne prennent leur sens que si l'on considère tout l'ensemble du développement dont ils font partie ; un même changement a une signification absolument différente suivant le procès dont il relève, et il n'est jamais légitime d'essayer d'expliquer un détail en dehors de la considération du système général de la langue où il apparaît.

Dès lors la nécessité s'impose de chercher à formuler les lois suivant lesquelles sont susceptibles de s'opérer les changements linguistiques. On déterminera ainsi, non plus des lois historiques, telles que sont les « lois phonétiques » ou les formules analogiques qui emplissent les manuels actuels de linguistique, mais des lois générales qui ne valent pas pour un seul moment du développement d'une langue, qui au contraire sont de tous les temps ; qui ne sont pas limitées à une langue donnée, qui au contraire s'étendent également à toutes les langues. Et, qu'on le remarque, ce ne seront ni des lois physiologiques ni des lois psychiques, mais des lois linguistiques. Quand on aura constaté par exemple que, entre deux voyelles, les consonnes tendent à subir certaines modifications, il faudra examiner si toutes les modifications observées se laissent ramener à une formule générale, en tant du moins qu'elles procèdent de la position intervocalique, et non de telle ou telle autre circonstance ; si l'on constate que ces altérations, dont l'aspect est au premier abord très divers, proviennent toutes d'une diminution de la force avec laquelle les consonnes sont prononcées, il ne restera plus qu'à rechercher la cause de cette faiblesse particulière, qui caractérise les consonnes articulées entre deux voyelles ; or, la cause ressort immédiatement de la formule même : les voyelles sont,

dans la syllabe, les éléments ouverts par excellence, ceux qui comportent le minimum d'articulation ; une consonne placée entre deux voyelles s'adapte aux éléments vocaliques précédents et suivants, tend à se vocaliser en quelque sorte, de la même manière qu'une voyelle placée près d'une nasale tend à se nasaliser.

De même l'élimination progressive des flexions complexes de l'indo-européen au cours du développement des divers dialectes se ramène à une formule compréhensive qui révèle la cause psychique du phénomène. Dès l'époque indo-européenne commune, les formes qui caractérisent une seule et même catégorie grammaticale variaient suivant les mots, et aussi en fonction d'autres catégories grammaticales ; elles variaient suivant les mots, ainsi l'optatif λύοι du présent grec λύω ne ressemble guère à l'optatif εἴη de εἰμι ; elles variaient suivant les catégories grammaticales, ainsi l'optatif présent λύοι de λύω est tout autrement constitué que l'optatif aoriste λύσειε du même verbe, et la première personne du pluriel εἶμεν « que nous soyons » diverge beaucoup d'avec la première du singulier εἴην. Or, partout et toujours, les langues tendent à abolir une pareille absence d'unité, et à instituer l'unité de forme pour l'unité de rôle grammatical et de signification. Ce résultat s'obtient par divers moyens, souvent par généralisation de l'un des procédés, ainsi quand la 1re personne du pluriel des verbes est caractérisée par *-ons* dans tous les verbes français. Un autre procédé très ordinaire est l'élimination des formes trop compliquées. Ainsi le français moderne a entièrement perdu le passé défini, du type *j'aimai, je fus, je dis*, dont les formations sont trop divergentes et la flexion trop différente de celle des autres formes verbales ; cette forme, trop riche en anomalies, ne survit plus que dans la manière de parler de certains Français qui subissent l'influence de patois locaux, et le français actuel ne connaît, en dehors de l'imparfait, d'autres formes du passé que celles qui sont constituées à l'aide d'un participe. Ce qui montre que pareille disparition n'a rien de fortuit, c'est qu'on en retrouve en slave le pendant exact ; le slave commun avait un aoriste qui présentait des caractères comparables

à ceux du passé défini français ; cet aoriste a été simplifié dans certains dialectes ; le russe et le polonais l'ont éliminé de très bonne heure, et les parlers serbes l'éliminent actuellement ; en slave comme en français, c'est une forme nominale qui tend à fournir l'unique expression du passé. Les langues de la Perse et de l'Inde présentent des développements tout pareils. Dans les mêmes conditions générales, on voit ainsi se réaliser les mêmes changements de formes grammaticales, et cela dans des conditions où tout soupçon d'influence mutuelle est naturellement exclu.

La recherche des lois générales, tant morphologiques que phonétiques, doit être désormais l'un des principaux objets de la linguistique. Mais, de par leur définition même, ces lois dépassent les limites des familles de langues ; elles s'appliquent à l'humanité entière.

Une famille de langues aussi grande, aussi variée, et qu'on peut suivre durant un aussi long espace de temps que la famille indo-européenne fournit assurément un champ d'observation assez vaste pour que les conclusions puissent prétendre à une valeur générale. En effet, il n'y a pas à objecter que l'expérience serait faussée par le fait même de la parenté des langues indo-européennes entre elles. L'affirmation de la parenté de deux langues n'implique la persistance d'aucun lien entre les deux ; elle suppose seulement un fait historique : des langues parentes sont des langues qui, à un certain moment du passé, n'en faisaient qu'une ; puis, à un moment ultérieur, des sujets parlant une langue ont été séparés par des circonstances quelconques de telle sorte qu'il y a eu deux groupes évoluant d'une manière indépendante. Il suit de cette définition que, du jour où la séparation des sujets parlants est accomplie, on se trouve en présence de deux développements distincts, et par là même de deux témoignages distincts, ayant chacun leur valeur propre dans les démonstrations de la linguistique générale. Durant les premiers temps après la séparation, l'identité de la langue parlée par les groupes examinés rend les conditions très semblables au point de vue proprement linguistique : l'expérience présente alors un

intérêt tout particulier à certains égards. Plus tard, quand les langues ont beaucoup divergé et que, comme c'est le cas actuel pour les langues de la famille indo-européenne, elles diffèrent profondément les unes des autres et n'ont presque plus rien de commun entre elles — quelle ressemblance y a-t-il entre l'anglais et le russe moderne ? — leur témoignage vaut en linguistique générale exactement ce que vaudrait le témoignage de langues de familles distinctes. Entre ces deux termes extrêmes, on a tous les degrés de dissemblance imaginables, et par là même une variété infinie d'expériences, dont la valeur en linguistique générale est inappréciable.

Néanmoins un sceptique pourrait tenter d'élever des doutes sur la portée des conclusions acquises, aussi longtemps qu'on n'aura pas suivi, pour autant que les faits attestés permettent de le faire, le développement de toutes les autres familles de langues. Les conditions différentes dans lesquelles se trouvent ces familles, la variété des combinaisons qu'elles offrent à l'étude permettront de vérifier la valeur des conclusions générales que l'étude des seules langues indo-européennes autorise à tirer ; elles permettront de plus de poser un certain nombre de questions qui, par des hasards divers, ne se posent pas aussi clairement dans les langues indo-européennes, et l'attention se trouvera ainsi attirée sur des détails peut-être importants, mais peu apparents, qui ont échappé jusqu'ici. La grammaire comparée des langues indo-européennes présente à ces recherches des modèles, et les méthodes qu'elle est parvenue à fixer éviteront aux travailleurs beaucoup d'inutiles tâtonnements. Déjà la grammaire comparée de certains groupes — le groupe sémitique, le groupe finno-ougrien par exemple — est très avancée ; ailleurs le travail est commencé et poussé assez avant pour qu'on en puisse déjà tirer profit, ainsi pour le turco-tatare, pour le bantou, pour le caucasique du sud, pour le malais. Au fur et à mesure que les grammaires comparées des divers groupes se constitueront d'une manière plus systématique, les lois de la linguistique générale acquerront plus de certitude, plus de précision et épuiseront plus complètement l'ensemble des faits de langue.

L'ancienne grammaire générale est tombée dans un juste décri parce qu'elle n'était qu'une application maladroite de la logique formelle à la linguistique où les catégories logiques n'ont rien à faire. La nouvelle linguistique générale, fondée sur l'étude précise et détaillée de toutes les langues à toutes les périodes de leur développement, enrichie des observations délicates et des mesures précises de l'anatomie et de la physiologie, éclairée par les théories objectives de la psychologie moderne, apporte un renouvellement complet des méthodes et des idées : aux faits historiques particuliers, elle superpose une doctrine d'ensemble, un système.

III

Toutes les lois générales qu'on a posées, toutes celles dont cette recherche, à peine entamée, réserve encore la découverte, ont cependant un défaut : elles énoncent des possibilités, non des nécessités.

Ainsi la loi relative à la débilité caractéristique de l'articulation des consonnes intervocaliques n'empêche pas les consonnes de subsister entre voyelles durant un temps illimité dans certaines langues. Le *t* intervocalique du mot indo-européen qui signifie « cent », celui du sanskrit *çatam* et du latin *centum*, subsiste dans le grec moderne *ekato* ; le *k* intervocalique du mot indo-européen pour « dix », celui du latin *decem*, subsiste sans la moindre altération dans le grec moderne *deka* ; de même, en lituanien, en slave, le *t* intervocalique du mot signifiant « mère » est aussi intact que celui du sanskrit védique *mātā*, du grec ancien μήτηρ, du latin *mater*, et l'on peut entendre un Lituanien dire *mote*, un Russe dire *mat'*, *māteri*, un Serbe dire *mati*, et ainsi dans tous les cas semblables. Toutes ces consonnes intervocaliques, toutes celles que présentent les exemples innombrables qu'il serait aisé d'énumérer, ont au moins quatre mille ans d'existence, et rien n'en fait prévoir la prochaine altération.

Les lois de la phonétique ou de la morphologie générale historique ne suffisent donc à expliquer aucun fait ; elles énoncent

des conditions constantes qui règlent le développement des faits linguistiques ; mais, même si l'on parvenait à les déterminer d'une manière complète et de tout point exacte, on ne saurait pour cela prévoir aucune évolution future, ce qui est la marque d'une connaissance incomplète ; car il resterait à découvrir les conditions variables qui permettent ou provoquent la réalisation des possibilités ainsi reconnues. Pour décisif que soit le progrès qui résulte de la constitution de la linguistique générale, on ne saurait donc s'en contenter.

L'élément variable qu'il reste à déterminer ne peut évidemment se rencontrer dans la structure anatomique des organes ou dans le fonctionnement de ces organes ; il ne se rencontre pas davantage dans le fonctionnement psychique : ce sont là des données constantes, qui sont partout sensiblement les mêmes, et qui ne renferment pas en elles des principes de variation. Mais il y a un élément dont les circonstances provoquent de perpétuelles variations, tantôt soudaines, et tantôt lentes, mais jamais entièrement interrompues : c'est la structure de la société.

Or, le langage est éminemment un fait social. On a souvent répété que les langues n'existent pas en dehors des sujets qui les parlent, et que par suite on n'est pas fondé à leur attribuer une existence autonome, un être propre. C'est une constatation évidente, mais sans portée, comme la plupart des propositions évidentes. Car si la réalité d'une langue n'est pas quelque chose de substantiel, elle n'en existe pas moins. Cette réalité est à la fois linguistique et sociale.

Elle est linguistique : car une langue constitue un système complexe de moyens d'expression, système où tout se tient et où une innovation individuelle ne peut que difficilement trouver place si, provenant d'un pur caprice, elle n'est pas exactement adaptée à ce système, c'est-à-dire si elle n'est pas en harmonie avec les règles générales de la langue.

A un autre égard, la réalité de la langue est sociale : elle résulte de ce qu'une langue appartient à un ensemble défini de sujets parlants, de ce qu'elle est le moyen de communication entre les membres d'un même groupe et de ce qu'il ne dépend

d'aucun des membres du groupe de la modifier ; la nécessité même d'être compris impose à tous les sujets le maintien de la plus grande identité possible dans les usages linguistiques ; le ridicule est la sanction immédiate de toutes les déviations individuelles, et, dans les sociétés civilisées modernes, on exclut de tous les principaux emplois par des examens ceux des citoyens qui ne savent pas se soumettre aux règles de langage, parfois assez arbitraires, qu'a une fois adoptées la communauté. Comme l'a très bien dit, dans son *Essai de sémantique*, M. Bréal, la limitation de la liberté qu'a chaque sujet de modifier son langage « tient au besoin d'être compris, c'est-à-dire qu'elle est de même sorte que les autres lois qui régissent notre vie sociale ».

Dès lors il est probable *a priori* que toute modification de la structure sociale se traduira par un changement des conditions dans lesquelles se développe le langage. Le langage est une institution ayant son autonomie ; il faut donc en déterminer les conditions générales de développement à un point de vue purement linguistique, et c'est l'objet de la linguistique générale ; il a ses conditions anatomiques, physiologiques et psychiques, et il relève de l'anatomie, de la physiologie et de la psychologie qui l'éclairent à beaucoup d'égards et dont la considération est nécessaire pour établir les lois de la linguistique générale ; mais du fait que le langage est une institution sociale, il résulte que la linguistique est une science sociale, et le seul élément variable auquel on puisse recourir pour rendre compte du changement linguistique est le changement social dont les variations du langage ne sont que les conséquences parfois immédiates et directes, et le plus souvent médiates et indirectes.

Il ne faut pas dire qu'on soit par là ramené à une conception historique, et qu'on retombe dans la simple considération des faits particuliers ; car s'il est vrai que la structure sociale est conditionnée par l'histoire, ce ne sont jamais les faits historiques eux-mêmes qui déterminent directement les changements linguistiques, et ce sont les changements de structure de la société qui seuls peuvent modifier les conditions d'existence du langage. Il faudra déterminer à quelle structure sociale répond une struc-

ture linguistique donnée et comment, d'une manière générale, les changements de structure sociale se traduisent par des changements de structure linguistique.

L'objet de ce cours sera donc de rechercher dans quelle mesure il est possible de reconnaître dès maintenant des rapports entre le développement linguistique et les autres faits sociaux. Les travaux préparatoires sont encore trop rares, les études de détail manquent encore trop pour qu'on puisse espérer donner des solutions définitives du premier coup. Mais il importe plus d'indiquer les problèmes nouveaux que pose le progrès de la science que de répéter les solutions, d'ailleurs nécessairement incomplètes, qu'ont reçues les vieux problèmes ; le devoir du professeur est, surtout ici, de montrer les recherches à entreprendre plus encore que de donner les résultats des travaux déjà faits. Le XIX^e^ siècle a été le siècle de l'histoire, et les progrès qu'a réalisés la linguistique en se plaçant au point de vue historique ont été admirables ; les sciences sociales se constituent maintenant, et la linguistique y doit prendre la place que sa nature lui assigne. Le moment est donc venu de marquer la position des problèmes linguistiques au point de vue social. Regarder vers l'avenir plutôt que vers le passé est le moyen de suivre l'exemple du maître qui m'a précédé dans cette chaire, et de demeurer fidèle à l'esprit de la noble maison qui m'a fait l'honneur de m'accueillir.

SUR LA MÉTHODE DE LA GRAMMAIRE COMPARÉE [1]

L'objet de la science qu'on est convenu d'appeler grammaire comparée est de faire l'histoire de développements linguistiques au moyen de rapprochements entre des langues diverses. Ce que doit déterminer le comparatiste dans chacun des cas qu'il étudie, c'est ce qui, parmi les faits examinés, suppose l'existence d'un seul et même idiome ancien, ce qui par suite résulte de la différenciation progressive d'une langue anciennement parlée d'une manière sensiblement une. Il doit faire un départ entre les rapprochements qu'il observe, et ne tenir compte que de ceux de ces rapprochements qui obligent à admettre un état de choses identique à un moment donné du passé dans le groupe des langues comparées.

Le problème de méthode consiste donc, étant donnés des faits linguistiques, à rechercher comment on peut reconnaître quels sont ceux de ces faits qui imposent, pour s'expliquer, l'hypothèse d'un point de départ identique.

On aperçoit immédiatement les deux types de questions qui se posent à tout historien : d'une part établir l'existence d'un certain fait ou d'un certain état de choses à un moment donné du passé, d'autre part poser le rapport qui existe entre faits de dates différentes. En l'espèce, les deux types de questions ne se laissent pas séparer.

Ce que l'on se propose ici, c'est de décrire le procédé de raisonnement des comparatistes et d'examiner quelle en est la valeur probante.

1. *Revue de métaphysique et de morale*, 1913, p. 1-15.

Le raisonnement est de la forme suivante : on observe dans telles et telles langues telles et telles manières de s'exprimer plus ou moins exactement concordantes ; ces concordances ne s'expliqueraient pas s'il n'y avait eu à une certaine date une forme commune dont toutes les formes semblables — mais en partie diverses — sont des continuations. Soit par exemple les pronoms personnels sujets dans les langues néo-latines :

	JE	TU
Roumain	*eo*	*tu*
Italien	*io*	*tu*
Vieux français	*jo*	*tu*
Espagnol	*yo*	*tu*

Il n'y a pas de raison générale pour que la personne qui parle se désigne par des formes telles que *eo*, *io*, *jo*, *yo* ; sans même s'adresser à des langues très différentes ou très éloignées, on voit que la personne qui parle se désigne tout autrement, en allemand, par *ich*, en russe, par *ja* (prononcer *ya*), en persan, par *man*. Il n'y a pas davantage de raison pour que la personne à qui l'on parle soit désignée par *tu* ; si les diverses langues indo-européennes ont conservé à cet égard des formes assez pareilles (allemand *du*, russe *ty*, etc.), il suffit de sortir du groupe indo-européen pour voir qu'on désigne la personne à qui l'on parle par *sen* en turc, par *anta* en arabe, par *o oe* en samoan, etc. Si donc, dans un groupe de langues, on a des formes concordantes pour désigner la personne qui parle et la personne à qui l'on parle, c'est que ces formes continuent une seule et même forme employée à date antérieure, et qu'il y a entre les deux une tradition, dont il reste à déterminer la nature. C'est précisément l'une des tâches principales de la linguistique historique que de déterminer les modalités diverses de la tradition linguistique ; les types généraux sont peu nombreux, mais le détail est d'une extrême complexité. On doit ici faire abstraction de ces faits particuliers, puisqu'il s'agit seulement de dégager le principe de la méthode.

Les raisonnements présentés ci-dessous sont de simples schèmes,

les choses sont en fait très complexes, et, comme en toute méthode historique, si le tact personnel et le bon sens ne guident pas constamment le chercheur, des erreurs énormes et ridicules se produisent à tout instant ; enfin, comme tout savant, le linguiste a sans cesse des vérifications de détail qui le dispensent de penser aux principes de la méthode. Les raisonnements indiqués ici ne représentent donc pas à proprement parler la manière de procéder des linguistes : on sait qu'il faut avoir pratiqué une science pour en saisir exactement la méthode et en contrôler les résultats. Seuls, ceux qui ont étudié les langues, savent combien sont divers les rapports de fait qui peuvent exister entre une langue et toutes celles dont on la rapproche pour en expliquer l'histoire.

Le point essentiel du raisonnement est celui-ci : la concordance de roumain *eo*, italien *io*, vieux français *jo*, espagnol *yo* ne peut pas être fortuite.

En effet les éléments phonétiques existant dans des langues comme l'italien, l'espagnol, le français, admettent un très grand nombre de combinaisons possibles. Si pour exprimer un même sens il se trouve que plusieurs langues recourent à une même combinaison (ou à des combinaisons qui s'expliquent par une même origine), cette rencontre exige une explication. Or, il n'y a dans les éléments phonétiques dont l'italien *io* est composé ou dans la manière dont ces éléments sont combinés rien qui explique l'usage fait de *io* pour désigner la personne qui parle. D'une manière générale, les tentatives qui ont été faites pour expliquer par des propriétés de la nature des sons le sens des mots n'ont jamais abouti à aucun succès. Le fait que les mêmes notions sont exprimées dans les diverses langues humaines par des sons infiniment divers et que le sens attribué aux mots varie sans que les sons y soient intéressés ou que, inversement, la prononciation des mots varie sans que le sens y soit intéressé suffit à montrer qu'il est inutile de rien chercher de ce côté.

D'autre part on sait par expérience qu'une même langue est parlée de manières sensiblement différentes par les divers individus qui l'emploient ; qu'elle est parlée de manières, encore beau-

coup plus différentes sur les divers points du territoire étendu où elle peut venir à être employée et aux époques successives où on la rencontre dans l'usage. Ces différences tendent constamment à augmenter. Si donc on rencontre chez deux groupes de sujets parlants des formes grammaticales identiques ou susceptibles d'être des transformations de formes identiques, on admettra que, à cet égard, ces deux groupes de sujets continuent la tradition d'une même manière de parler.

Il n'y a pas lieu d'essayer de chiffrer, au moyen du calcul des probabilités, quelles chances il y a pour qu'une notion donnée soit exprimée par une combinaison de sons donnée. Les notions à exprimer sont choses trop variables et trop peu précises pour se laisser saisir ; les combinaisons de sons ne sont pas toutes également admises par une langue donnée ou par les langues en général. D'autre part certaines combinaisons phoniques s'associent mieux à certaines notions que certains autres sons, et ces associations ont facilité la fixation de ces groupements, pour exprimer ces notions : les verbes *craquer* ou *glisser* ont une forme phonétique qui s'associe dans notre esprit au sens qu'ils expriment ; mais ceci n'est pas nécessaire ; et des sons analogues expriment des notions tout autres : *traquer* ou *braquer* n'indiquent pas des bruits, *plisser* n'indique pas un mouvement aisé et continu. Les conditions qui entrent en jeu sont trop multiples, elles échappent trop à une appréciation numérique pour que l'on puisse faire intervenir un calcul.

Sans recourir au calcul, qui n'a rien à faire ici, on a le droit d'affirmer qu'une rencontre de quatre langues dans l'expression de la première personne par des formes aussi pareilles, du type *yo*, peut malaisément être tenue pour un accident. Et si, comme il arrive en l'espèce, toute une série d'autres langues présentent des formes analogues : *eu* en roumain, *iou* en rhéto-roman, etc., le hasard devient plus invraisemblable.

Le hasard est plus exclu encore si l'on observe que la personne à qui l'on parle est désignée en italien et en espagnol par *tu* (prononcé *tou*), en français par *tu* ; que la personne qui parle se désigne avec une ou plusieurs autres, en italien par *noi*, en espa-

gnol par *nos*, en français par *nous*; qu'une forme de pronom de la première personne servant de complément est en italien *mi*, en espagnol *me* (prononcé *mé*), en français *me*, etc.

Dans le cas des langues citées, italien, espagnol, français, rhéto-roman, etc., les concordances observées s'expliquent aisément : ces langues sont celles de populations qui occupent des pays ayant tous fait pendant plusieurs siècles partie de l'empire romain, dont la langue était le latin. Or, le latin est attesté par de nombreux textes et bien connu. Toutes les concordances signalées s'expliquent par le fait que les formes en question continuent des formes latines : *ego* « moi », *tu* « toi », *nos* « nous », *uos* « vous », *me* « me », etc. La démonstration de l'origine commune des formes considérées qui a été obtenue par la comparaison trouve dans des données historiques une vérification. C'est une heureuse rencontre, qui a facilité l'établissement de la théorie, mais qui n'ajoute, au fond, rien à la preuve. Car, d'une part, la démonstration est suffisante sans cela et, de l'autre, rien n'empêche que là où se parlait le latin durant la période de l'empire romain, il se parle aujourd'hui une langue qui ne soit pas néo-latine; ainsi en Illyrie, où le latin a été, sous l'empire romain, la langue courante et où il a subsisté longtemps des parlers néo-latins, on n'emploie plus aujourd'hui que des parlers slaves; le français et l'italien parlés actuellement en Tunisie ne sont pas des continuations du latin qui s'y parlait sous l'empire romain. La preuve des parentés de langues fournie par la méthode comparative se suffit à elle-même, et elle est la seule valable.

La preuve résulte tout entière de ce qu'il est improbable que certaines concordances existent là où il n'y a pas une tradition historique commune. Tout le problème se ramène donc à celui-ci : quels sont les types de concordances qui supposent cette communauté de tradition.

Les concordances qui reconnaissent des causes générales, valables pour l'ensemble des langues, sont dénuées de valeur probante pour le comparatiste historien. Par exemple la présence d'un type de consonnes qui se trouve presque partout, comme le *t* et le *k*, n'établit à aucun degré que deux langues soient

parentes ; l'emploi universel de *t* ou de *k* tient à ce que la structure des organes articulatoires rend aisée la réalisation de ces types phonétiques. Au contraire, la présence dans deux langues voisines d'un même phonème de type tout particulier et rare parmi les langues humaines, comme le *jery* du russe et du polonais, est une première raison de croire à la parenté de ces deux langues ; ce n'est d'ailleurs qu'une simple indication, et ce n'est pas sur un phénomène de ce genre, de nature encore beaucoup trop générale, qu'on peut fonder l'affirmation d'une parenté.

Ce qui établit une origine commune, c'est l'existence concordante dans deux ou plusieurs langues de particularités telles qu'elles ne s'expliquent pas par des conditions générales, anatomiques, physiologiques ou psychiques. Le fait que tous les pronoms envisagés ci-dessus sont des mots courts ne prouve rien ; car il tient au rôle que les pronoms jouent dans la phrase, et presque jamais, dans aucune langue, les pronoms ne sont des mots de plus de deux syllabes ; le plus souvent, ils sont monosyllabiques. Ce qui tend à prouver une parenté, c'est que la première personne du singulier soit caractérisée par un *y* ou par un phonème qui, comme le *j* français, s'explique aisément par un ancien *y*, et par une voyelle du timbre *o* ou capable de sortir aisément de *o* ; c'est que la deuxième personne du singulier soit caractérisée par un *t* et par un *u* (*ou* français) ou par une voyelle susceptible de sortir aisément d'un ancien *u* (*ou*), comme l'*u* français ou le *jery* russe ; c'est que la première personne du pluriel soit caractérisée par un *n* suivi d'un *o* (ou d'une voyelle qui s'explique par l'existence d'un ancien *o*) et d'un autre élément ; et ainsi de suite, car il n'y a pas de raison générale pour que la première personne du singulier soit caractérisée par *y* plutôt que par *t*, par *n* ou par *v*, et pas de raison pour que la voyelle qui suit la consonne soit celle qui se rencontre dans chacune des formes citées, et non une autre.

Du principe de la méthode il résulte que les faits probants en matière de grammaire comparée sont des faits particuliers, et ils sont d'autant plus probants que, par leur nature, ils sont moins suspects de pouvoir reconnaître une cause générale. Il n'y a rien

là que de naturel : puisqu'il s'agit de poser par des procédés comparatifs le fait historique de l'existence d'une langue particulière, c'est-à-dire une chose qui, par définition, se produit en vertu d'un concours de circonstances diverses n'ayant pas de rapports nécessaires les unes avec les autres, ce sont des faits particuliers de caractère historique qui doivent seuls entrer en considération.

La preuve est particulièrement nette là où l'on observe des variations concomitantes ; ainsi la troisième personne du singulier du verbe « être » est de la forme *è* en italien, *est* en français, *es* en espagnol ; la troisième personne du pluriel, de la forme *sono* en italien, *sont* en français, *son* en espagnol. Cette opposition d'une forme reposant sur un ancien *est* et d'une forme reposant sur un ancien *sont* (latin *sunt*) est propre aux langues néo-latines. D'autres langues indo-européennes offrent des faits parallèles : l'allemand par exemple a *ist* et *sind*. En dehors des langues indo-européennes, cette manière d'exprimer le verbe « être » par une racine *es-* alternant avec *s-* et une désinence *-ti* au singulier, *-enti* ou *-onti* au pluriel ne se rencontre jamais. Dans le parallélisme particulier des faits néo-latins relatifs à *il est*, *ils sont*, et dans le parallélisme un peu moins complet des faits latins, germaniques, slaves, etc., relatifs à ce même verbe, on a une preuve de la parenté spéciale des langues néo-latines entre elles et de la parenté plus lointaine du latin, du germanique, du slave, etc., langues qui toutes sont des formes prises par un même idiome, l'indo-européen.

Car un état de choses aussi singulier résulte du concours de toute une série de circonstances : l'expression du verbe « être » par *es-*, de la troisième personne du singulier par *-ti*, et de la troisième du pluriel par *-enti* ou *-onti*, l'emploi d'une alternance de forme radicale avec un *e* au singulier et sans *e* au pluriel ; il est improbable qu'aucune de ces circonstances considérée à part vienne à se reproduire exactement dans deux langues d'une manière indépendante, plus improbable encore que ces quatre circonstances se retrouvent simultanément à un moment donné d'une manière indépendante dans deux langues différentes. Deux

langues qui présentent de pareilles concordances sont donc des formes diverses prises par une seule et même langue, en l'espèce l'indo-européen. En toute science, démontrer, c'est découvrir des faits qui apparaissent « évidents » ; on est ici devant les « évidences » de la grammaire comparée : certaines concordances de faits particuliers sont telles qu'il est « évident » qu'elles ne sauraient se trouver dans deux langues différentes ; si donc les deux langues considérées sont distinctes — et souvent au point que les sujets parlants ne s'entendent pas entre eux —, c'est que ces deux langues sont deux aspects pris par une seule et même langue en différents lieux et en différents temps.

C'est par des faits particuliers de ce genre qu'on établit les parentés de langues. Là où, comme il arrive souvent, la structure des langues considérées ne fournit que peu ou ne fournit pas de faits *singuliers* qui puissent être rapprochés, l'établissement rigoureux de parentés de langues rencontre les plus graves difficultés, et la linguistique historique arrive à peine à se constituer. Au contraire, là où, comme sur le domaine indo-européen, sur le domaine sémitique, sur le domaine finno-ougrien, sur le domaine bantou, sur le domaine indonésien (malais), les concordances *singulières* de cet ordre abondent, la linguistique historique est déjà créée et progresse rapidement.

Des analogies de structure, même grandes, si elles ne sont pas accompagnées de faits particuliers significatifs, ne prouvent pas une parenté de langues.

On a souvent parlé d'une grande famille de langues ouralo-altaïques, comprenant à la fois le finno-ougrien (groupe du finnois et du magyar, etc.) avec le samoyède, le turc et le mongol et même le japonais. En effet il y a entre toutes ces langues des ressemblances frappantes de structure générale. Mais, aussi longtemps qu'on n'aura pas reconnu des faits particuliers communs, comme ceux à l'aide desquels on a établi l'existence d'un groupe finno-ougrien et samoyède, on n'aura pas le droit de parler d'une famille de langues ouralo-altaïques, c'est-à-dire d'affirmer que le finno-ougrien (finnois, magyar, etc.) et le turc par exemple sont des transformations diverses d'une seule et même

langue ayant existé à un moment du passé. Telle langue indo-européenne a pu prendre un aspect général pareil à celui qu'offrent ces langues : l'arménien moderne a une structure grammaticale analogue à celle du turc, et pourtant l'origine de la grammaire de l'arménien moderne est indépendante du turc. Et l'on ne rapproche pas l'arménien moderne du turc parce que le détail des formes qui indiquent le nombre et le cas des noms, le nombre, la personne et le temps des verbes est entièrement différent en arménien et en turc.

Pour donner une idée plus précise de l'application des principes précédents, il ne sera pas inutile d'indiquer avec quelque détail comment on peut établir l'étymologie d'un mot.

Faire l'étymologie d'un mot, c'est déterminer toute l'histoire de la manière dont ce mot a été transmis dans une langue donnée depuis une date donnée.

Par exemple ce mot peut s'être transmis de génération en génération entre les deux dates considérées ; et alors il peut avoir subi des changements spéciaux tenant à des conditions particulières : le sens peut être demeuré le même, ou il peut avoir varié, soit que les objets désignés se soient modifiés avec le temps, soit que les groupes sociaux qui emploient le mot ordinairement viennent à changer de nature ou d'étendue, soit enfin que le mot ait reçu dans la phrase des rôles nouveaux. Ou bien le mot peut être entré dans la langue à un moment compris entre les deux dates considérées ; et alors il peut avoir été créé de toutes pièces — cas très rare semble-t-il —, ou formé à l'aide d'éléments existant dans la langue, ou emprunté à un autre idiome, proche ou lointain. On voit que les problèmes qui se posent, dont il n'est possible de donner ici qu'un aperçu général, sont infiniment nombreux si l'on tient compte du fait que les conditions dont il vient d'être question sont sujettes à changer constamment et qu'un mot peut, dans un bref espace de temps, être soumis à l'action de conditions très diverses. Ce n'est pas poser au linguiste un problème simple que de lui demander l'étymologie d'un mot, c'est lui demander de décrire une histoire complexe, et dont il est impossible de prévoir le degré de complexité.

Même les cas qui paraissent au premier abord le plus simples sont encore très compliqués. Par exemple, le mot français *père* est l'un des mots, assez peu nombreux, qui semblent s'être transmis d'une manière continue de génération en génération depuis l'époque de la langue indo-européenne commune jusqu'au latin classique et jusqu'au français d'aujourd'hui. Mais il a subi dans sa forme des modifications profondes : entre la façon dont le mot se prononçait en indo-européen et celle dont il se prononce aujourd'hui en français il n'y a qu'un trait commun, le fait que le mot commence par un *p* ; la forme grammaticale a changé du tout au tout, puisque le mot indo-européen admettait plus de quinze formes différentes suivant les nombres et les cas et que le mot français est invariable (l'*s* du pluriel est purement orthographique). Le sens n'a pas moins changé ; le mot indo-européen désignait un rôle social ; le « père » était le chef d'une maison, et c'est pour cela que le mot a une valeur religieuse ; il sert à désigner le « ciel père », le *Jupiter* romain ; encore en latin le mot *pater* a un sens avant tout social ; ce mot désigne une situation de famille, non un fait naturel ; et le procréateur est désigné plutôt par un autre mot, qui a expressément ce sens, *genitor* ; actuellement, au contraire, le « père » est celui qui est physiologiquement le procréateur, et ceci est si vrai que la langue populaire emploie pour les animaux le mot *père* au sens de « mâle » et dit d'un lapin que c'est un *père* pour indiquer que c'est un mâle ; ou bien *père* désigne tout homme d'un certain âge, non pas avec la nuance de respect qu'un Romain mettait quand il appelait un homme âgé un *pater*, le traitant par là de *chef*, mais au contraire de manière familière et peu respectueuse ; ou, si le mot *père* est employé avec une nuance de respect, c'est, dans la langue spéciale de l'église catholique, pour s'adresser à des moines qui, de par leur profession, ne sont jamais des chefs de famille. On voit combien, dans un cas pourtant très favorable à la stabilité du mot considéré, les changements de toutes sortes sont profonds de l'indo-européen au français, c'est-à-dire durant un laps de temps qu'on peut évaluer à quelque quatre mille ans.

En général, le vocabulaire évolue vite. Des mots dont la valeur

s'affaiblit rapidement par l'usage sont renouvelés pour obtenir une expression plus intense. Des changements dans les choses, des changements sociaux, des usages qui empêchent momentanément ou pour toujours l'usage de certains vocables, les événements historiques, des emprunts, souvent innombrables, à des langues étrangères ou à des parlers de groupes sociaux particuliers, et, d'autre part, les changements de prononciation, les innovations qui s'introduisent dans la grammaire, une infinité de circonstances variées font que le vocabulaire de la plupart des langues change sans cesse : depuis le XVII^e^ siècle, la prononciation et la grammaire du français employé par la bonne société parisienne n'ont changé que dans le détail ; il ne faudrait pas un grand effort à un contemporain de Molière pour s'adapter au système linguistique du français d'aujourd'hui ; mais il comprendrait très mal ce qu'on lui dirait parce que les mots et le sens des mots lui seraient en grande partie nouveaux. Et quiconque n'a pas fait de la langue du XVII^e^ siècle une étude approfondie est exposé tantôt à ne pas comprendre un texte de cette époque et tantôt à faire en le lisant de grossiers contresens, simplement parce que les mots employés sont, pour une large part, autres que les nôtres, ou qu'ils ont un sens autre.

Il résulte de cette variabilité du vocabulaire que donner l'étymologie des mots est chose difficile, souvent impossible pour peu qu'il s'agisse d'une langue dont on connaît mal l'histoire. Pour pouvoir affirmer qu'un mot est la continuation de tel ou tel mot connu à une date antérieure ou qu'il a été formé à l'aide de tels ou tels éléments, il faut, suivant le principe posé ci-dessus, déterminer des concordances de détail précises dont la rencontre ne puisse passer pour fortuite. Une longue expérience a maintenant appris aux linguistes que les ressemblances de son ou de forme qu'on constate au premier abord sont souvent trompeuses, et ce n'est plus de ressemblances vagues et indéfinies que s'autorisent les étymologistes méthodiques.

En ce qui concerne la forme, les linguistes s'attachent à déterminer des correspondances régulières. Soit par exemple deux mots de forme analogue dans les langues néo-latines.

	PÈRE	MÈRE
Italien	*padre*	*madre*
Espagnol	*padre*	*madre*
Provençal	*paire*	*maire*
Français	*père*	*mère*

De pareils parallélismes montrent que les ressemblances de forme extérieure que l'on observe entre les mots de chaque série s'expriment par des formules précises et que par suite elles ont chance de n'être pas accidentelles. C'est uniquement à l'aide de ces formules, jamais à l'aide de ressemblances apparentes, qu'un linguiste fait des étymologies. Tant qu'on n'est pas arrivé à poser des formules de correspondances de cette sorte, on tâtonne. Pour les langues où les formules sont reconnues, quiconque essaie de faire des étymologies sans connaître et sans appliquer les formules fait une besogne vaine.

Un mot est défini par l'association d'un sens donné à un ensemble donné de sons susceptible d'un emploi grammatical donné. Pour avoir une valeur, une concordance entre deux mots doit donc porter à la fois sur les sons, sur le sens et, s'il y a lieu, sur l'emploi grammatical. Plus la concordance est parfaite *à la fois* aux trois points de vue et plus l'étymologie a de chances d'être correcte.

Les concordances entre les sons doivent être établies au moyen des formules de correspondances dont il vient d'être question. Elles ont d'autant plus de valeur qu'elles s'étendent à un nombre d'éléments phonétiques plus grand. Une concordance portant sur quatre ou cinq ou six éléments a beaucoup moins de chances d'être fortuite qu'une concordance portant sur un ou deux éléments seulement. Il est donc très difficile de faire des étymologies sûres dans les langues où les mots sont courts, bornés à une consonne suivie d'une voyelle comme il arrive souvent. Un mot comme latin *pater, patris* qui comprend quatre éléments caractéristiques (*p, a, t, r*) ou comme *frāter*, qui en contient cinq (*f, r, a, t, r*) fournit un point de départ excellent à la recherche. Au contraire les mots monosyllabiques à voyelle

finale, qu'on rencontre souvent dans les langues du Soudan ou de l'Extrême-Orient ne se prêtent que mal à l'étymologie. Les combinaisons du type de *ba* ou *ta*, qu'on peut former avec deux éléments, sont trop peu nombreuses pour prouver beaucoup si on les rencontre dans deux langues à la fois.

A l'égard du sens, les concordances ont d'autant plus de valeur que les mots considérés ont des sens plus précis et rigoureusement déterminés. Si la notion d'un même animal est attachée à un groupe nombreux d'éléments phonétiques se répondant suivant les formules de correspondances établies comme italien *cavallo*, provençal *cavall*, français *cheval*, espagnol *caballo*, pareille concordance exclut évidemment le hasard. Plus les sons considérés sont vagues et plus l'étymologie devient incertaine; car la concordance entre certains sons et une certaine notion devient moins frappante. Les étymologies de « racines » à sens vague et très général sont beaucoup plus incertaines que les étymologies de mots bien définis, sauf bien entendu au cas où l'on peut mettre en évidence l'existence dans les langues rapprochées de mots bien définis à sens bien défini appartenant à une racine : il n'y a pas de rapprochement plus sûr que celui de grec φέρω, latin *fero*, gotique *baira* « je porte », arménien *berem* « je porte », sanskrit *bharāmi* « je porte », c'est-à-dire le présent de la racine *bher-* dans les diverses langues indo-européennes. Il arrive que les mots changent de sens : la différence de sens entre les mots rapprochés doit être expliquée par des raisons précises, autant que possible par des raisons tirées de faits positivement attestés, car il n'existe pas de règles sémantiques permettant de dire que tel ou tel développement de sens est exclu.

Les explications de noms propres auxquelles on se complaît souvent et dont beaucoup de linguistes aiment à tirer des conclusions historiques ont peu de valeur. La force probante d'une étymologie provient de ce que l'on ne peut tenir pour fortuit que dans deux ou plusieurs langues un même sens soit exprimé par des sons identiques ou susceptibles d'être ramenés à une identité antérieure en vertu des formules de correspondances ; mais si l'on rapproche un nom propre d'une langue d'un nom

commun d'une autre en soutenant que le nom commun fournit une interprétation du nom propre il est clair que toute la force de la preuve disparaît, puisque le sens attribué au nom propre est arbitraire. On ne peut donc interpréter des noms propres que là où l'explication est évidente; par exemple il y a en français une série de noms propres qui sont des noms de professions : *Charpentier*, *Bouvier*, *Vacher*, *Pasteur*, *Boucher*, *Boulanger*, etc. ; il est donc légitime de penser que *Sueur* est le mot ancien français *sueur* « cordonnier » (lat. *sutor*) et *Fèvre* l'ancien nom du « forgeron » (lat. *faber*). Si l'on avait quelques doutes, on serait rassuré par le fait que *Lesueur* est à *Sueur* ce que *Lecomte* est à *Comte*, *Lepelletier* à *Pelletier*. Hors les cas évidents de cette sorte, les interprétations de noms propres ne sauraient être prouvées et n'ont, par suite, que la valeur de jeux d'esprit.

Les rapprochements reçoivent des confirmations utiles quand on peut constater que des concordances grammaticales s'ajoutent à la concordance du son et du sens. Si, par exemple, on pouvait douter de la concordance de l'italien *cuocere* et du français *cuire* garantie par l'exacte concordance des sons (malgré la première apparence) et du sens, la concordance des participes *cotto* et *cuit* (*cuite*) ferait beaucoup pour lever le doute. Les langues qui, comme les langues indo-européennes (surtout dans la période ancienne) et les langues sémitiques, ont des particularités grammaticales attachées à certains mots se prêtent donc mieux à la démonstration de l'étymologie que les langues où tous les mots se conforment aux mêmes règles grammaticales. La difficulté qu'on éprouve à poser la grammaire comparée de certaines langues, notamment en Extrême-Orient, vient en partie de là.

Ce qui rend l'étymologie encore plus incertaine et difficile qu'il ne paraît à en juger par ce qui vient d'être dit, c'est que les mots qu'on trouve dans une langue donnée proviennent toujours de plusieurs langues diverses et que, d'autre part, des mots nouveaux sont sans cesse formés, à l'aide soit d'éléments existant dans la langue, soit d'éléments empruntés à d'autres langues. On n'a donc jamais le droit d'affirmer que tel ou tel mot ne peut s'expliquer qu'à l'aide de tels ou tels éléments existant à un

moment donné dans telle ou telle langue. Les difficultés de fait que rencontre l'étymologiste sont innombrables, et, s'il est aisé d'apercevoir des ressemblances entre les mots, il faut être du métier pour savoir à quel point il est malaisé d'établir la justesse d'un rapprochement étymologique.

Des faits cités il ressort que le seul moyen de preuve employé dans les démonstrations de grammaire comparée consiste à constater des concordances entre certains sons (ou certains procédés d'expression très définis) et certaines notions, et à examiner si ces concordances peuvent être fortuites. Le degré de valeur de la preuve tient uniquement au degré de probabilité que la concordance ne puisse pas être accidentelle. De là vient que les étymologies sont probables à des degrés très divers ; certaines, comme le rapprochement du français *père* et de l'italien *padre*, de *cuire* et d'italien *cuocere*, peuvent passer pour sûres ; mais d'autres rapprochements sont simplement possibles, et entre une probabilité qui touche à la certitude et une simple possibilité, il y a tous les degrés de probabilité imaginables.

On peut utiliser la méthode comparative d'une autre manière et se demander ce qu'il y a de commun à toutes les langues humaines ou à plusieurs langues indépendamment du fait que certaines de ces langues sont les transformations d'une même langue ayant existé antérieurement, comme le français et l'italien sont des transformations du latin, comme l'arabe de Syrie, l'arabe d'Égypte, et l'arabe du Maroc — très différents aujourd'hui dans l'usage parlé — sont des transformations de l'arabe des conquérants arabes. On aboutirait ainsi à constituer une linguistique générale. Tel n'est pas l'objet de ce qu'on appelle « grammaire comparée ». La linguistique générale qu'on obtient en faisant abstraction de l'histoire est une science encore peu faite, difficile à faire et qui pour se faire suppose du reste qu'on ait déjà décrit aussi complètement que possible l'histoire du plus grand nombre de langues possible. Ce qui sert à l'établissement de la linguistique générale est ce qui est dénué de valeur pour la grammaire comparée et inversement. Par exemple le fait que

le sémitique et l'indo-européen ont également une distinction nette du nom et du verbe ne prouve pas que ces deux langues soient apparentées ; mais le fait que deux langues nettement distinctes à l'époque historique ont, par des procédés différents, distingué le nom du verbe est l'une des preuves qu'on peut alléguer à l'appui de la doctrine générale que toutes les langues tendent à distinguer ces deux catégories grammaticales essentielles. — Le sémitique avait une déclinaison à trois cas, l'indo-européen une déclinaison à huit cas ; or, on observe que, au cours du développement historique de la langue sémitique et de la langue indo-européenne, le nombre de ces cas tend à se réduire et que chaque nom tend à avoir une seule forme ou, tout au plus, deux formes distinctes, l'une pour le singulier, l'autre pour le pluriel ; ce parallélisme de développement ne prouve aucune parenté ; il s'explique par des causes générales qui ont agi séparément dans chacune des langues sémitiques et dans chacune des langues indo-européennes, et qui ont d'ailleurs agi à des degrés très différents ; car, si l'anglais, le persan, le français par exemple ont perdu toute flexion casuelle, comme le syriaque ou l'arabe moderne, c'est par suite d'évolutions relativement récentes dont on suit l'histoire. Une concordance générale de développement ne peut servir à démontrer une communauté d'origine historique. Le comparatiste historien doit se méfier avant tout de ce que les biologistes appellent des phénomènes de convergence.

Ce qui vient d'être dit de la linguistique s'applique au fond à toute méthode comparative. Il y a deux manières d'employer la comparaison : ou bien l'on constate des concordances particulières qui ne pourraient être fortuites, et l'on en conclut à une origine historique commune. Ou bien l'on constate des concordances entre tous les faits examinés, et l'on aboutit à poser des lois générales. Le second procédé qui, en des matières aussi complexes que les faits sociaux, présente d'extrêmes difficultés n'a pas à être discuté ici. Le seul but qu'on s'est proposé était de faire ressortir comment avec une méthode comparative on peut conclure à des faits de caractère historique. En linguistique, la méthode a fait ses preuves, et de nombreuses confirmations de

fait en ont établi la légitimité. Ailleurs, notamment en matière de religion, d'institutions politiques ou juridiques, de technique industrielle, elle réussit moins: nulle part, autant qu'en matière de langue, les procédés employés n'admettent des combinaisons multiples; nulle part par suite, les combinaisons réalisées n'ont aussi peu de chances de se retrouver fortuitement. Quand on voudra se rendre compte des erreurs que commettent souvent les linguistes ou les personnes qui se mêlent de raisonner de linguistique historique, il suffira de remonter au principe de la démonstration; du même coup on sera en garde contre les erreurs que l'on est exposé à commettre toutes les fois que l'on emploie la méthode comparative à reconstituer des évolutions de faits sociaux autres que le langage, et l'on verra dans quelle mesure il est légitime d'employer, en un cas donné, un pareil moyen de preuve.

NOTE

SUR

UNE DIFFICULTÉ GÉNÉRALE DE LA GRAMMAIRE COMPARÉE

A M. MICHEL BRÉAL

A L'OCCASION

DU 25e ANNIVERSAIRE DE SON ÉLECTION A L'INSTITUT

(3 décembre 1875 — 3 décembre 1900).

I

Si l'on envisage d'une part la structure grammaticale de l'indo-européen telle que l'examen des langues les plus anciennement attestées de la famille, surtout du sanskrit védique et du grec homérique, permet de la définir, et si l'on passe en revue d'autre part les grammaires des langues indo-européennes à date plus récente, on constate que le plan général de la langue a été changé dans toutes d'une manière essentielle. Toutes ont perdu les mêmes traits de la grammaire en les remplaçant par d'autres qui se retrouvent partout, non pas identiques, mais sensiblement analogues : l'histoire des prépositions en est un bel exemple (Bréal, *Sémantique*, 203). En un mot, les langues indo-européennes se sont sur bien des points développées indépendamment les unes des autres d'une manière *parallèle*.

Il n'y a là rien de surprenant. Car si peu que l'on sache sur les causes des changements linguistiques, il est au moins aisé de déterminer trois ordres de faits dont ils dépendent : la structure de la langue considérée, les conditions générales physiologiques et psychiques de l'existence du langage, et enfin les influences particulières qui s'exercent en un temps et en un lieu donnés. Les langues indo-européennes ont divergé parce qu'elles ont subi chacune certaines actions extérieures auxquelles les autres

ont échappé. Mais elles se sont à beaucoup d'égards développées parallèlement, parce que, le point de départ étant le même pour toutes, les possibilités ou même les nécessités de changements qui résultaient de la structure de l'indo-européen et des conditions générales d'évolution du langage étaient identiques. En observant ces développements parallèles, on a donc un moyen de déterminer, en même temps que les conditions générales du développement linguistique, ce qui est essentiel à l'indo-européen et, du même coup, ce qui, étant particulier à chaque dialecte et provenant d'influences propres à ce dialecte, le caractérise.

En fait, aussi longtemps que des accidents historiques graves n'interviennent pas, les divers dialectes d'une même langue évoluent d'une manière parallèle, *et cela sans qu'il y ait lieu d'admettre aucune influence des uns sur les autres*. Les langues slaves fournissent de ce fait capital toute une série d'exemples. Pour ne rien dire ici du plus net de tous, le remplacement de la 1re personne verbale *dělajǫ* par *dělam*, on peut citer la flexion de l'adjectif. Le slave commun avait deux formes d'adjectifs, l'une simple, fléchie comme le substantif, qui était employée avec valeur indéterminée, l'autre composée de l'adjectif simple et d'un démonstratif enclitique, qui était employée avec valeur déterminée; par suite de diverses transformations, les unes phonétiques, les autres morphologiques, le caractère composé de la forme déterminée a cessé d'être sensible dès le début de la période historique du slave, et les dialectes slaves se sont trouvés posséder deux flexions de l'adjectif, l'une identique à celle des substantifs, l'autre analogue à celle des démonstratifs ; or, dans tous ces dialectes, la forme déterminée a tendu de bonne heure à perdre son sens propre, ce qui s'explique en grande partie par le fait que, le slave n'exprimant pas la détermination là où il n'y a pas d'adjectif, l'expression de la détermination dans l'adjectif allait contre le système général de la langue; et peu à peu la forme déterminée, analogue à la flexion des démonstratifs et qui par là même permettait de distinguer l'adjectif du substantif (distinction que l'indo-européen ne paraît pas avoir faite dans la forme, mais qu'on en est venu par la suite à marquer de plus en plus

presque partout), est devenue la flexion normale de l'adjectif, tandis que la forme simple tendait à disparaître ; partout la forme simple a laissé des traces, mais partout aussi la forme composée a empiété sur l'autre. Si donc, par un hasard, les textes anciens des divers dialectes slaves où les formes simples sont régulièrement employées avaient disparu, on serait certainement amené par la comparaison des dialectes modernes à exagérer beaucoup l'importance de la forme composée en slave commun aux dépens de la forme simple. — Par suite de l'extension de la forme déterminée, les cas obliques du pluriel de l'adjectif se trouvent n'avoir plus de distinction de genre et c'est ce qui a provoqué sans doute au masculin et au neutre des substantifs l'emploi de la finale *-ami* de l'instrumental féminin pluriel : ce fait s'est produit à date historique et en partie récente dans plusieurs langues slaves bien distinctes et tout à fait séparées les unes des autres.

Pour voir que les divers dialectes indo-européens ont subi avant d'être écrits des modifications parallèles de même ordre que les modifications ainsi constatées en slave à l'époque historique, il suffit d'envisager dans ses traits les plus généraux le système morphologique indo-européen. Trois procédés d'expression morphologique étaient employés simultanément pour caractériser chaque forme : 1° la division en racine, suffixe et désinence ; 2° le vocalisme ; 3° le ton. La racine jouait dans la morphologie indo-européenne un rôle analogue à celui qu'elle joue dans l'arabe du Coran et non moins important ; elle était une réalité d'autant plus saisissable pour l'esprit que, la plupart du temps, elle suffisait déjà par elle-même à constituer un thème sans l'adjonction d'aucun suffixe et que les thèmes racines nominaux ou verbaux étaient régulièrement en usage, comme le sanskrit védique permet encore de l'entrevoir. Du fait même que les trois éléments : racine, suffixe et désinence, étaient bien distincts pour l'esprit résultait la possibilité pour chacun d'avoir une voyelle caractéristique : et en effet, un mot indo-européen n'est pas défini si l'on sait seulement quel suffixe est ajouté à sa racine, quelle désinence à son thème ; on doit savoir de plus quel est le voca-

lisme de la syllabe prédésinentielle et de la syllabe présuffixale. Enfin pour achever la détermination morphologique du mot indo-européen il faut savoir s'il a un ton et, au cas où il en a un, sur lequel des éléments morphologiques le ton est placé ; on sait que le ton consistait en une simple montée de la voix et que par suite la syllabe qui en était frappée n'acquérait pas par là une importance démesurée comme la syllabe frappée de l'accent allemand ou de l'accent russe ; l'équilibre des parties du mot n'était en rien compromis par la présence du ton sur l'une d'elles, et par suite le ton ne troublait pas la clarté de la division morphologique du mot, ni ne détournait l'attention de la valeur significative de chaque voyelle : à ce point de vue le mot indo-européen différait radicalement du mot de toute langue moderne. Le datif védique *mánase* par exemple est défini complètement si l'on sait : 1° que la racine est *man-*, le suffixe *-as-*, la désinence *-e* ; 2° que la présuffixale et la prédésinentielle ont toutes deux le vocalisme *a* (indo-européen *e*) ; 3° que le ton est sur la racine ; mais il n'est défini que si l'on a les trois données.

Dans tous les dialectes, ce système si cohérent a été disloqué, et l'on en ignorerait l'existence si par bonheur le véda n'en avait conservé les traits essentiels et si l'on n'en retrouvait par ailleurs assez de restes pour établir qu'il a été le système indo-européen commun et que chaque idiome l'a brisé à sa manière. Chacune des trois parties qui le constituaient a été altérée indépendamment et, en même temps, par son altération même, a contribué à l'altération des deux autres. La racine a cessé d'être nettement perceptible du jour où les thèmes-racines n'ont plus été employés d'une manière fréquente ; le suffixe a cessé d'être bien isolé de la désinence par suite de l'extension toujours croissante des formes thématiques aux dépens des formes athématiques et de l'introduction des voyelles de liaison ; dès lors le vocalisme de la présuffixale et de la prédésinentielle et la place du ton (sur la racine, le suffixe ou la désinence) perdaient leur signification caractéristique. Le vocalisme a été fortement obscurci par des changements phonétiques, par ceux notamment relatifs aux sonantes *y*, *w*, *r*, *l*, *m*, *n*, dont les unes se sont définitivement

rapprochées des voyelles et les autres des consonnes ; par là même l'indépendance de chacun des éléments morphologiques du mot disparaissait et le ton était détaché de *la* voyelle de l'élément morphologique à laquelle il appartenait d'abord essentiellement. Enfin le ton a disparu, comme en indo-iranien ou en germanique, ou est devenu accent d'intensité comme on l'observe, à une époque historique, en russe : dans le second cas, non seulement son rôle morphologique se réduisait, mais un système complet d'alternances vocaliques s'étendant à toutes les syllabes, comme en indo-européen, devenait impossible ; car, seule, la syllabe accentuée attire l'attention.

Il serait très long de mettre en évidence toutes ces actions et réactions dont la conséquence est claire : nulle part la morphologie indo-européenne n'a subsisté, et là même où, comme dans l'Inde, au début de la tradition, on la retrouve sinon intacte du moins conservée, à ce qu'il semble, dans ses traits essentiels, elle n'a pas tardé à s'altérer aussi et plus profondément qu'ailleurs. Toutes les langues indo-européennes ont perdu les mêmes procédés grammaticaux, et si, comme en germanique, il reste quelque chose des alternances vocaliques, ou, comme en russe, quelque chose à la fois des alternances vocaliques et des mouvements du ton, ce ne sont que d'informes débris : la morphologie régulière n'en conserve plus rien.

La flexion indo-européenne était extrêmement complexe, et beaucoup plus qu'on n'est porté à le croire au premier abord. La flexion védique même ne donne sans doute qu'une idée très atténuée de cette complexité. La tendance de toutes les langues a été de réduire cette richesse ancienne de la flexion à une simplicité relative ; ainsi, au lieu de la multiplicité des formes verbales védiques ou homériques, une conjugaison à deux thèmes seulement s'est établie presque partout à date plus ou moins ancienne : par exemple en latin où tout verbe a deux thèmes l'un de présent, l'autre de prétérit fournissant chacun une série de formes (*facio, faciebam, faciam*, etc., et *feci, feceram, fecero*, etc.). Ces deux thèmes ne sont pas partout les mêmes, et dans chaque langue ils fournissent des « temps » et des modes différents. Mais

on peut poser presque en règle universelle du développement du verbe indo-européen que, parti d'un système complexe de formations indépendantes les unes des autres toutes rattachées à une racine, il a tendu vers une conjugaison rigide comportant seulement deux thèmes distincts.

Cette tendance à la simplification de la flexion ne se manifeste pas seulement dans les changements essentiels de tout le plan de la grammaire, elle apparaît aussi dans une foule de détails. Par exemple le duel qui en indo-européen constituait une troisième série de formes à côté de celles du singulier et du pluriel se conserve tant que la flexion reste complexe, ainsi en védique, et disparaît quand la flexion se simplifie, ainsi dans les prākrits; il n'est pas fortuit que le slave qui a maintenu à peu de chose près toute la richesse en cas de la déclinaison indo-européenne emploie encore le duel au x^e^ siècle après J.-C., tandis que le grec qui, dès l'époque historique, a réduit sa déclinaison à quatre formes casuelles au singulier et au pluriel tend à le perdre de bonne heure; déjà dans la langue homérique le duel n'est plus employé avec constance, l'ionien ni l'éolien n'en ont plus trace, et l'attique même, qui l'a mieux conservé, le perd tout à fait de 360 à 320 avant J.-C. Les raisons de la disparition du duel indiquées par M. Delbrück, *Vergleichende Syntax*, III, p. 251 et suiv., sont visiblement insuffisantes : cette disparition suppose l'action d'une cause générale, à savoir la grande tendance à la simplification de la flexion.

On expliquerait de même que les deux modes à signification voisine, le subjonctif et l'optatif, ne coexistent plus que dans les textes anciens des deux langues connues à la date la plus ancienne : l'indo-iranien et le grec : partout ailleurs il ne reste qu'une forme, issue soit du subjonctif, soit de l'optatif. — Il serait aisé de multiplier les exemples de ces réductions grâce auxquelles les morphologies récentes ont atteint le degré de simplicité et de régularité qu'on observe notamment en gotique et en arménien.

L'évolution des verbes à nasale fournit un exemple plus remarquable peut-être que tous les autres. M. F. de Saussure a reconnu, dans son admirable et célèbre *Mémoire*, que le type

indo-iranien des verbes tels que skr. *riṇákti*, zend *irinaxti* remonte à l'indo-européen et a expliqué par là le vocalisme et la formation de sanskrit *pṛṇáti*, *punáti*, etc., en regard de *pūrṇá-*, *pūtá-*, etc.; on explique de même skr. *vṛṇóti* en regard de arménien *gelum* gr. ϝελύ-σθη, lat. *uoluo*. Partant de cette découverte, M. Pedersen a montré (*Indogermanische Forschungen*, II, 285 et suiv.) comment plusieurs types de verbes à nasale qu'on rencontre dans les diverses langues sont issus du type à infixe. Le type à infixe n'a subsisté partiellement qu'en baltique (où il a pris un développement et a fourni des dénominatifs tels que lituanien *bąlù* « je blanchis, je deviens blanc ») et en latin, mais en devenant thématique, conformément à une tendance générale de la flexion indo-européenne : on a vieux prussien *-linka* et lat. *linquit* en regard de skr. *riṇák-ti*, pluriel *riñcánti*; il s'agit ici d'un développement récent, car de même que de la racine **bheid-* attestée par gotique *beitan* et peut-être par gr. φείδομαι, le latin a *findit* en regard de skr. *bhinátti*, plur. *bhind-ánti*, le prākrit a *bhindadi*. Conformément au plan général de la langue qui, usant fréquemment des suffixes, ne connaissait nul infixe autre que celui-ci, le type à infixe a été éliminé partout ailleurs de diverses manières; le plus souvent un suffixe a été substitué à l'infixe, c'est ce qui est arrivé régulièrement en slave où *bůnǫ*, *svĭnǫ*, etc. répondent à lituanien *bundù*, *szvintù*, etc. et en arménien où *lkhanem*, *gtanem*, etc. répondent à skr. *riṇákti*, *vindáti* (gâthique *vīnasti*), etc.; et, dans plus d'un cas, on ne saurait démontrer l'existence d'un ancien infixe que par la divergence même des types, ainsi dans arménien *lizanem* « je lèche » ou gr. λιχνεύω, en face de lat. *lingō*. Les verbes en **-nā-* et en **-neu-* qui, comme l'a si bien vu F. de Saussure, ne s'expliquent que par l'hypothèse d'un ancien infixe, ont été pris pour des verbes à suffixe et par suite ont fourni des verbes nouveaux à suffixe : ainsi l'ancienne forme à infixe attestée par skr. *yunákti*, lat. *iungō*, lit. *jùngiu* est remplacée en grec par ζεύγνυμι dont le caractère récent est dénoncé par le vocalisme. Il a parfois fallu très peu de chose pour transformer en type à suffixe un ancien type à infixe : par exemple de la racine **k*^{w}*reyə-* « acheter » (gr. πρίασθαι)

on a un verbe à infixe *$k^{w}rinā$- attesté par pāli *kiṇāti*, vieux russe *kŕĭnuti* (*krenut'*), v. irlandais *crenim*, gallois *prynu* ; il a suffi au sanskrit classique de substituer à la brève de **kriṇāti* la longue de *krītá*- et de dire *krīṇāti* pour donner à *-nā*- le caractère d'un suffixe régulier dans ce verbe et dans les autres de forme analogue. Nulle part mieux que dans cette histoire des verbes à nasale on ne voit à quel point les dialectes historiquement connus se sont écartés du type indo-européen, perdant tous les mêmes particularités caractéristiques, et remplaçant d'anciens types par des procédés conformes au plan général de la langue.

De cet exemple ressort clairement une difficulté fondamentale de la grammaire comparée : les ressemblances que présentent les langues indo-européennes entre elles et sur lesquelles seules peut reposer une théorie de l'indo-européen admettent souvent deux interprétations : identité initiale ou développement dialectal identique : dans les deux cas les formes observées dans les diverses langues font au premier abord l'effet de reposer sur un état premier un. La question qui se pose est alors de déterminer laquelle des deux interprétations est la vraie. Là où apparaissent des divergences comme dans les verbes à nasale, on ne peut pas ne pas voir que les dialectes indo-européens ont évolué parallèlement de très bonne heure. Mais dans les cas où l'évolution a été sensiblement identique, le résultat est le même que s'il y avait eu unité dès le début : ainsi en slave quand on constate partout à l'époque moderne l'extension de la forme composée de l'adjectif ; alors on ne saurait rien décider. Parfois le changement, tout en portant dans les divers dialectes sur les mêmes points de la grammaire, s'est produit dans chacun de manière différente et alors il est impossible de faire une restitution quelconque, à moins qu'un dialecte n'ait à peu près conservé l'état ancien, comme l'a fait le sanskrit pour les verbes à nasale : l'obscurité impossible à dissiper de toute une partie des formes de la déclinaison indo-européenne provient d'un fait de ce genre. Il y a là, pour la grammaire comparée, une difficulté fondamentale qui tient à l'essence même de sa méthode et qu'on n'a pas le droit de jamais perdre de vue.

LINGUISTIQUE HISTORIQUE

ET

LINGUISTIQUE GÉNÉRALE [1]

La grammaire se fait de deux manières : elle est ou descriptive, ou historique. Dans le premier type, on se borne à exposer l'usage linguistique d'un groupe d'hommes donné, à un moment donné, en un lieu donné. Dans le second, on expose les changements de l'usage entre deux époques plus ou moins éloignées l'une de l'autre.

Pour procéder avec rigueur en grammaire historique, il faudrait comparer l'usage exactement au même lieu et chez les seuls descendants des hommes qui constituaient le groupe parlant la langue à la première époque considérée ; mais les changements historiques et sociaux sont en général trop rapides et trop fréquents pour que cette rigueur soit réalisable ; en fait, on est amené à comparer des états de langues à des dates différentes sur des domaines sensiblement différents, les sujets parlants étant des hommes d'origines très diverses, le plus souvent mal connues ou ignorées.

La grammaire descriptive et la grammaire historique ne diffèrent pas essentiellement l'une de l'autre. D'une part, en effet, toute description est en quelque mesure historique : si un que soit le groupe social où une langue est parlée, les divers sujets qui le composent sont, à certains égards, à des degrés différents de l'évolution qui emporte constamment chaque langue : chaque génération nouvelle apporte quelques menues innovations, si

1. *Scientia* (*Rivista di scienza*), vol. IV (1918), nº VIII.

bien que le parler des vieillards diffère souvent d'une manière sensible de celui des jeunes gens. De plus il peut y avoir dans le groupe des éléments conservateurs qui maintiennent les archaïsmes, et des éléments novateurs où au contraire l'évolution est en avance. Enfin, il est des usages qui tendent à devenir de plus en plus rares tandis que d'autres n'apparaissent d'abord qu'à l'état de tentatives isolées. Toute description précise et complète d'une situation linguistique à un moment donné comporte donc la considération d'une certaine part d'évolution ; et ceci est inévitable puisque une langue qui se parle n'est plus par là même en état de stabilité complète.

D'autre part, les moyens dont dispose la grammaire historique ne permettent jamais de décrire d'une manière vraiment continue la courbe suivie par l'évolution. Pour des faits contemporains, certains observateurs ont eu occasion d'examiner — et encore par une méthode un peu indirecte — la façon dont quelques changements se sont produits. M. l'abbé Rousselot a pu suivre ainsi la transformation de *l* mouillée en *y* dans un patois français ; M. Gauchat a fait des observations analogues sur un patois suisse roman. Toutefois ce sont des observations isolées, précieuses en ceci qu'elles permettent de se former une idée de la façon dont les langues évoluent, mais dont on n'a jamais l'équivalent pour les périodes plus anciennes, et dont on a même très peu d'exemples encore à l'époque actuelle. En fait et pratiquement, la grammaire dite historique consiste à juxtaposer des grammaires descriptives de plusieurs époques successives et à constater qu'à un fait *a* d'une première époque répond un fait *b* d'une seconde, un fait *c* d'une troisième, et ainsi de suite. Il est aisé de le montrer.

Ce que l'on appelle une « loi phonétique » exprime purement et simplement une correspondance entre deux états linguistiques donnés. Cette correspondance peut avoir lieu entre deux états d'une même langue à deux époques distinctes. Ainsi toutes les langues romanes sont des formes plus récentes du latin. Dire que *f* du latin devient *h* en espagnol moderne, c'est dire qu'à latin *f*, répond espagnol *h*, ainsi *hija* à *filiam*. Quand on a plusieurs

transformations différentes d'une même langue antérieure, on observe de même des lois de correspondance entre langues considérées à un même moment ; en ce sens l'espagnol *hija* répond à italien *figlia*, portugais *filha*, français *fille*, roumain *fie-* (dans *fiema*), etc., c'est-à-dire que espagnol *h* répond à *f* des principales langues romanes modernes. Tant à l'un qu'à l'autre point de vue, on ne constate rien autre qu'un fait de correspondance. Seulement, dans le premier cas, la loi de correspondance exprime une transformation qui s'est opérée à un certain moment entre la date où est attestée la forme latine *filiam* et celle où est attestée la forme espagnole *hija*. Dans l'autre cas au contraire, le sens de la correspondance est au premier abord incertain ; il faut un examen pour pouvoir décider de quel côté il y a innovation. Si même on n'avait pas le latin, on arriverait du reste à conclure sans hésitation que l'espagnol a innové ; ce qui le prouve, ce n'est pas que la plupart des langues romanes aient *f* en regard du *h* espagnol et gascon (le gascon concorde ici avec l'espagnol), car il arrive qu'une seule langue conserve un état ancien altéré dans toutes les autres. On notera cependant que la concordance exacte de toutes les langues autres que l'espagnol et le gascon (qui a *hil'* en face de l'espagnol *hijo* « fils ») s'expliquerait malaisément si *f* résultait d'une innovation : l'innovation ne serait sans doute pas exactement identique dans toutes les langues. Toutefois, il n'y a là qu'une probabilité ; la vraie preuve est ailleurs, dans un fait de théorie générale : *f* peut se transformer en *h* sans aucune influence externe ; il suffit pour cela que le mouvement de rapprochement de la lèvre inférieure avec les dents d'en haut soit exécuté incomplètement ; au contraire il ne semble pas que *h* puisse se transformer spontanément en *f*, et nulle part en effet on n'observe un changement spontané de *h* en *f*.

C'est donc une règle de phonétique générale qui, dans le cas examiné, permet de décider d'un fait historique. La chose est du reste confirmée en espagnol même : tandis que en gascon, le passage de *f* à *h* est général et qu'il ne subsiste d'*f* en aucun cas, l'espagnol maintient *f* en certaines situations, notamment devant *ue* et représente le latin *focum* par *fuego* ; l'espagnol *fuego* répond

donc à italien *fuoco*, portugais *fogo*, français *feu*, roumain *foc*, etc., alors que le gascon a *huek*.

Quoi qu'il en soit des faits particuliers, le linguiste n'envisage dans tous ces cas que des correspondances ; ce sont les seuls faits positifs qui lui soient donnés ; le reste n'est que théorie et construction hypothétique. Ces correspondances supposent des changements réels qui ont eu lieu à certaines dates en certains lieux ; mais elles n'indiquent rien sur les faits réels dont elles résultent ; on ne peut sur ces faits faire que des suppositions plus ou moins plausibles, et ceci même pour des développements voisins de l'époque actuelle et pour les idiomes les mieux connus.

On a envisagé ici le cas des langues romanes où l'on dispose de la forme ancienne qui a évolué de manières diverses : le latin. Là où l'on ne possède pas la forme ancienne, on n'a, à plus forte raison, affaire qu'à des systèmes de correspondances. Soit par exemple le démonstratif sanskrit *sā*, gotique *sō*, grec (dorien) *hā*, (ionien-attique) *hē* « celle-ci » ; on y observe une correspondance : sanskrit et gotique *s-*, grec et iranien *h-* ; c'est le seul fait positif dont on dispose, et il résulte de la simple confrontation des quatre systèmes considérés. Si l'on possédait la langue dont le sanskrit, le grec, le gotique et l'iranien sont les représentants historiquement attestés, on constaterait à coup sûr que la forme représentée par sanskrit *sā*, etc. avait *s* et non pas *h* : car *s* est sujet à se transformer en *h*, mais l'inverse n'est pas vrai ; cette conclusion ne repose que sur une doctrine de linguistique générale.

Ce qui vient d'être dit des correspondances phonétiques s'applique aussi aux correspondances de formes grammaticales. Le procédé par lequel les formes grammaticales se transforment ne diffère de celui par lequel elles se reproduisent que dans l'application : il s'agit toujours de l'analogie. L'analogie qui fait faire la forme régulière *vous téléphonez* sur *téléphoner*, *nous téléphonons*, etc., d'après le modèle de *laisser*, *nous laissons*, *vous laissez*, ne diffère pas de celle qui conduit à une forme contraire à l'usage ancien dans *vous disez* (au lieu de *vous dites*), sur *dire*, *nous disons*, d'après le modèle de *lire*, *nous lisons*, *vous lisez*, et des verbes pareils.

Il n'y a ainsi au fond, en ce qui concerne l'étude positive des langues particulières, qu'une seule discipline grammaticale, à la fois descriptive et historique, et qui met seulement en évidence le côté descriptif ou le côté historique suivant le but spécial de la recherche entreprise. Et qu'elle soit plutôt descriptive ou plutôt historique, l'étude n'a pour objet que des faits particuliers. Car, soit qu'on la considère en un temps et un lieu donnés, soit qu'on en suive le développement en des lieux et en des temps divers, une langue n'est qu'un fait particulier. Ainsi conçue, l'étude n'aboutit donc, comme toute autre étude historique, qu'à des conclusions particulières.

Mais les personnes qui ont poussé un peu avant leurs études linguistiques ont acquis bien vite le sentiment que, dominant ces faits particuliers propres à chaque langue, il existe des principes qu'il faudrait dégager pour diriger et assurer les recherches. Et l'on n'envisage pas ici les sciences dont relève indirectement la linguistique, parce que les faits dont elle s'occupe sont soumis à leurs lois ; il va de soi que la linguistique est dominée par les principes de l'acoustique, de l'anatomie et de la physiologie en tant qu'il s'agit de l'émission et de l'audition des sons, de la psychologie, en tant qu'il s'agit de la perception de la parole, de la sociologie, en tant qu'il s'agit de l'action des parlers individuels les uns sur les autres et de l'adaptation des langues à la communication entre les hommes. Il a été fait un grand effort (notamment depuis l'apparition de la *Sprache* de M. Wundt), pour faire ainsi profiter la linguistique des acquisitions des sciences voisines, et il est résulté de là un progrès considérable dont les effets se font et se feront sentir de plus en plus. Ce n'est pas ce progrès qu'on se propose d'examiner dans le présent article, uniquement consacré à la linguistique au sens étroit du mot.

Sans sortir de la discipline grammaticale proprement dite, il semble qu'il soit possible de dégager des principes. Sans doute ces principes devront s'expliquer en dernière analyse par les conditions physiques, anatomiques, physiologiques, psychiques, sociales dans lesquelles se trouvent les sujets parlants. Mais ils

ont un caractère purement grammatical et sont par là même de nature à diriger les recherches grammaticales. Pour les dégager, il faudrait un livre qui n'est pas encore fait et qui n'est sans doute pas encore assez préparé par des recherches de détail pour être écrit dès maintenant. Mais il a déjà été assez publié sur la question (M. Grammont est le savant qui en a le premier pris une conscience claire et qui a le mieux et le plus nettement posé ce genre de problèmes) pour qu'on puisse marquer au moins le caractère des résultats qu'on doit attendre sur ce domaine.

Les principes linguistiques dont il s'agit ici ne sauraient naturellement avoir un caractère de nécessité. En effet la seule nécessité qui se rencontre en linguistique est celle qui s'impose aux sujets parlants d'un même groupe de conserver un même système linguistique afin de se comprendre aisément entre eux. Cette obligation assure le maintien du système, ou quand il intervient des changements, l'identité de ces changements à l'intérieur d'un même groupe linguistique. Chaque système linguistique est d'ailleurs si cohérent que, à l'intérieur d'un groupe donné où les conditions sont à peu près semblables, les changements ne sauraient différer beaucoup. Dans la mesure où il se produit des changements purement individuels, ils s'éliminent sans difficulté, au plus tard à la mort de l'individu qui a innové. Mais cette nécessité, qui résulte du besoin qu'ont les hommes de s'entendre entre eux, ne porte que sur les moyens d'expression particuliers à chaque langue, et ne s'étend précisément pas aux principes.

Ces principes exprimeront donc simplement les possibilités que peuvent utiliser les langues particulières ; certaines possibilités sont réalisées d'une manière nécessaire dans chaque langue, mais sans qu'on puisse jamais prévoir *a priori* lesquelles. On s'est déjà servi ci-dessus de deux principes élémentaires de ce genre, à savoir que *ʃ* peut devenir *h*, et *s* devenir *h*, mais que les transformations inverses n'ont pas lieu spontanément ; là où l'on observe le passage de *h* à *f* dans les parlers arméniens par exemple il y a eu influence de phonèmes voisins, et le passage de *h* à *s* n'est sans doute attesté nulle part. La linguistique

historique est dominée par des principes de cette sorte, et les linguistes s'en servent constamment, mais presque toujours sans les formuler d'une manière explicite, et souvent sans en avoir nettement conscience. Dans la mesure où ils valent, ces principes sont d'ailleurs précis, comme le montrent les exemples cités, et ils déterminent les limites entre lesquelles peut se mouvoir le changement linguistique. Il importe de les formuler, ne fût-ce que pour en examiner la portée et pour savoir dans quelle mesure on peut faire fond sur eux. C'est un travail qui est entamé, mais dont une petite partie seulement est accomplie. On voudrait essayer de marquer par quelques exemples les résultats qu'on a le droit d'en attendre. Pour simplifier, les exemples seront tous empruntés à la phonétique ; mais il existe des principes dominants en morphologie aussi bien qu'en phonétique.

La structure phonétique du langage articulé est partout la même dans ses traits essentiels. Partout le discours se compose de sons tenus qu'on appelle *voyelles* séparés par des phonèmes qui en général se rapprochent plus du type des bruits et qui sont d'ordinaire caractérisés par un mouvement plus ou moins prononcé de fermeture de l'un des organes articulatoires suivi d'un mouvement d'ouverture ; ce sont ces articulations qu'on nomme *consonnes*. L'ensemble d'une tenue avec le mouvement d'ouverture qui la commence et le mouvement de fermeture qui la termine constitue la *syllabe*. Définir exactement la syllabe est chose difficile, et qui a donné lieu à beaucoup de discussions. La structure de la syllabe varie d'ailleurs sensiblement d'une langue à l'autre. Mais dans toutes les langues le discours se répartit ainsi en groupes élémentaires assez brefs consistant en tenues de sons séparées par des mouvements articulatoires plus ou moins étendus.

On montrerait de même que toutes les voyelles se classent entre la plus ouverte *a* et les plus fermées, *i* et *u*, ces dernières pouvant jouer le rôle de consonnes si elles se trouvent près de tenues ; on montrerait aussi que les consonnes se répartissent entre les occlusives, qui comportent un arrêt total du passage de l'air suivi d'explosion, et les diverses sortes de continues, qui

comportent seulement un resserrement plus ou moins prononcé du passage de l'air en un point quelconque des organes articulatoires.

Les occlusives ne sont jamais que de trois sortes : les *gutturales*, dans lesquelles l'occlusion est réalisée par le dos de la langue appuyé contre le palais (le terme de *gutturale* est donc absurde); les *dentales*, dans lesquelles l'occlusion est réalisée par le bord externe de la langue appuyé contre le palais ; les *labiales*, où l'occlusion résulte du rapprochement des lèvres. Il est rare qu'aucun de ces trois types d'occlusives manque totalement dans une langue. Pareille classification ne signifie pas qu'une langue ne puisse avoir que trois sortes d'occlusives, car il existe bien des sortes de gutturales, et bien des sortes de dentales ; le sanskrit par exemple a deux séries dentales qui sont encore conservées dans les langues modernes de l'Inde : les dentales proprement dites, pour lesquelles le bord externe de la langue est appuyé à la hauteur des alvéoles, et les *cérébrales* pour lesquelles le bord de la langue est appuyé plus haut, contre le palais, sensiblement au-dessus des alvéoles. Parmi les langues modernes, le français, l'italien et l'allemand par exemple ont des *dentales* du type sanskrit, et l'anglais des *cérébrales*, ce qui fait que les Hindous actuels transcrivent par la notation de leurs *cérébrales* le *t* et le *d* de l'anglais.

On peut donc établir une phonologie générale, c'est-à-dire déterminer quels sont les types articulatoires existant dans le langage. Il va de soi que cette doctrine générale ne permettra jamais de prévoir ce que sera une articulation de l'un des types dans une langue donnée. D'une part, en effet, plusieurs des types comportent des variétés distinctes, et dont les différences peuvent être très grandes ; on vient de voir ainsi qu'il y a plusieurs sortes de dentales suivant la place où le bord de la langue vient s'appliquer. D'autre part, un même phonème comporte plusieurs mouvements articulatoires dont chacun a ses limites de jeu propres. Ainsi les occlusives peuvent être ou non accompagnées du resserrement de la glotte qui produit les vibrations glottales, et elles peuvent être prononcées avec une application plus ou

moins ferme de l'organe d'occlusion ; enfin elles peuvent être à peu près fixes quelle que soit la voyelle suivante, ou elles peuvent être très diverses suivant la nature de cette voyelle : des occlusives dont l'occlusion caractéristique a lieu en un même point peuvent donc différer profondément. Ainsi le *t* français et le *t* russe sont assez pareils en ce qui concerne le début des vibrations glottales, le degré de force et le point d'occlusion ; mais ils diffèrent beaucoup en ce qui concerne l'influence de la voyelle suivante. C'est au contraire par le degré de force de l'occlusion et par le moment où commencent les vibrations glottales que le *t* allemand faible et « aspiré », se distingue du *t* français fort et non aspiré. Le *t* italien se distingue du *t* français par la netteté de son occlusion et par le silence bien perceptible qui suit cette occlusion. La connaissance du type général ne dispense donc jamais d'une description précise de chaque fait particulier ; mais elle donne le moyen de situer chaque particularité dans un ensemble.

Le degré de stabilité d'un phonème fait partie de sa description. Il y a des phonèmes qui sont bien adaptés à nos organes, qui se retrouvent par suite dans un grand nombre de langues, et qui, une fois établis, ne disparaissent que difficilement. C'est ainsi que la sifflante *s* est un phonème largement répandu et que possèdent beaucoup d'idiomes. Il y a au contraire des phonèmes rares, qui n'existent que dans peu de langues et qui ne survivent guère à certaines conditions spéciales ; ainsi l'ensemble du slave a eu à date ancienne une voyelle nommée *jery*, qui existe encore en russe et en polonais, mais qui est devenue *i* dans les autres idiomes slaves, parce que les circonstances qui conditionnent le maintien de la prononciation du *jery* n'ont subsisté qu'en russe et en polonais.

Cette remarque générale éclaire des faits particuliers qui apparaissent dans l'histoire de certaines langues et qui ont longtemps paru mystérieux. En voici un exemple.

Quand un type articulatoire est atteint par un changement à un certain point de vue, tous les phonèmes du même type doivent être touchés par le même changement, sinon en même

temps, du moins dans des périodes de temps rapprochées ; et si l'on envisage deux termes suffisamment éloignés, les correspondances doivent être parallèles. Ainsi à toutes les occlusives sourdes de la plupart des autres langues indo-européennes : *p*, *t*, *k* du latin, du grec, du slave, du sanskrit, etc., le germanique commun répond par les spirantes *f*, *th* (*th* anglais), *x* (*ch* allemand). Or, on observe des cas où, isolément, indépendamment des autres occlusives, une spirante *f* répond à une occlusive *p*. Dans les mêmes langues, l'occlusive sonore *b*, l'occlusive nasale *m* se maintiennent souvent, attestant clairement que l'articulation labiale n'est pas devenue impossible. Ainsi en éthiopien et en arabe, c'est *f* qui répond au *p* des autres langues sémitiques, alors que l'éthiopien et l'arabe ont *t* et *k* comme toutes les autres langues de la famille ; les mêmes langues ont *b* et *m* comme le reste du sémitique. Cette correspondance fait au premier abord l'effet d'un étrange caprice. Elle cesse d'apparaître comme un hasard si l'on constate que des langues très diverses, sur les points les plus éloignés du monde, en offrent de pareilles. En ossète (dialecte iranien parlé dans quelques vallées du Caucase), le *p* iranien commun est représenté par *f*, qui est bien conservé à l'initiale du mot ; le fait semble très ancien dans le groupe dialectal iranien dont l'ossète est le dernier survivant ; les noms iraniens conservés dans les inscriptions grecques des rivages septentrionaux de la mer Noire présentent en effet un φ (*f*) grec en regard du *p* des autres dialectes iraniens, par exemple *furto-* en regard de *puθra-* « fils » du reste de l'iranien (voir V. Miller, *Sprache der Osseten*, p. 6 et 33). Le nicobar répond par *f* à un *p* des autres dialectes du groupe mon-khmer (et, d'une manière plus générale du groupe malayo-polynésien qui appartient à la même famille) ; le nicobar a par exemple *fōan* « quatre » en regard de mon *pan*, khmer *puon*, mais *kōan* « enfant » tout comme mon *kon*, khmer *kūn* (voir P. W. Schmidt, *Die Mon-Khmer Völker*, p. 85 et suiv.). M. Carlos Everett Conant, dans une note (f *and* v *in Philippine languages*, p. 3 et suiv.), signale des parlers des Philippines qui répondent par *f* au *p* malais. Le *p* finno-ougrien est représenté en magyar par *f*, ainsi magyar

fa « arbre » en face de finnois *puu*, votyak et ziriane *pu*, etc. (v. Szinnyei, *Finnisch-ugrische Sprachwissenschaft*, p. 24). Si le japonais n'a pas de *p* simple dans ses mots indigènes, c'est que, comme l'indiquent divers indices, le *p* y a passé à *f*, et celle-ci à *h* : le *h* japonais se trouve ainsi alterner en certaines conditions soit avec la sonore *b*, soit avec la géminée *pp* ; et *h* occupe dans l'alphabet japonais la place du *p* (voir Kanazawa, *Ueber den Einfluss des Sanskrits auf das japanische Schriftsystem*, Tokio, 1907, p. 18). En arménien, où le *t* et le *k* de l'indo-européen oriental sont rendus régulièrement par *th* et *kh*, le *p* initial de mot donne, non pas le *ph* qu'on attend, mais *h* (issu évidemment d'une *f* plus ancienne ; car, à l'intérieur du mot, entre voyelles, *p* est représenté par *w*, *v*) ; ainsi l'on a *hayr* « père » en face du latin *pater*. Ces faits permettent d'interpréter le celtique où l'évolution est plus avancée : au *p* indo-européen le celtique répond par l'absence de toute consonne ; l'irlandais a par exemple *athir* « père » en regard du latin *pater* ; on voit sans peine que *p* est devenu *f*, celle-ci *h*, et qu'enfin *h* a cessé d'être prononcé ; la spirante *f* a laissé une trace devant *t*, parce qu'alors elle a passé à la spirante gutturale *ch* : « sept » se dit en irlandais *secht*, ce qui répond au latin *septem*[1]. On sait donc maintenant que *p* est sujet à passer à *f* dans des conditions pareilles à celles où *k* et *t* se maintiennent. Le changement a lieu régulièrement dans un certain nombre de langues, où tous les *p* sont alors atteints ; on peut donc toujours s'attendre à ce que ce changement se réalise ; mais ce n'est qu'une possibilité ; et, en fait, la grande majorité des langues où *t* et *k* subsistent conservent aussi *p* inaltéré. Même sans connaître la raison pour laquelle *p* est plus sujet à devenir une spirante que ne le sont *t* et *k*, on a ainsi une donnée précieuse. Cette raison se laisse d'ailleurs entrevoir : l'occlusive *p* n'a pas pour se réaliser un point d'appui solide tel qu'est le palais contre lequel la masse musculaire de la langue vient s'appuyer ; la lèvre d'en bas vient s'appuyer à la masse musculaire de la lèvre d'en

1. Pour un fait analogue dans une langue américaine, voir Sapir, *Journal de la Société des Américanistes*, XI (1919), p. 453.

haut. Ce qui montre que là est la difficulté, c'est que *p* est altéré sans que *b* et *m* le soient : l'occlusion difficile à réaliser est l'occlusion relativement forte de la sourde *p* ; les occlusions moins fermes de *b* et de *m* ne font pas autant de difficulté.

C'est pour la même raison que la spirante bilabiale *f*, réalisée par le rapprochement des lèvres, est instable. Ou bien elle passe à *h* comme on l'a vu dans les exemples qui viennent d'être cités du japonais, de l'arménien, du celtique (car c'est à *f* bilabiale que passe le *p* qui devient spirant) ; ou bien elle devient labio-dentale, et dès lors elle est stable. Le passage de *f* bilabiale à *f* labio-dentale et de *w* bilabial à *v* labio-dental est un fait que l'on observe fréquemment et dont on peut citer un nombre illimité d'exemples (latin, allemand, etc.). Et, du jour où la prononciation labio-dentale, avec la lèvre d'en bas appuyée contre les dents d'en haut, est le type normal, on n'a presque plus de changements à attendre ; car ce type de phonèmes est aisé à prononcer et est trop distinct de tous les autres pour être exposé à beaucoup de changements.

Les faits précédents ont montré quelles différences de stabilité on peut observer entre les phonèmes. Il est plus important encore de rechercher comment un même mouvement articulatoire est traité suivant les conditions où il doit être exécuté. L'écriture nous a en effet habitués à considérer comme identiques des articulations qui ont un trait dominant commun mais qui diffèrent à beaucoup d'égards. Ainsi le *p* de *père* et celui de *repère* s'écrivent par la même lettre *p*, et l'effet acoustique est à peu près le même en effet dans les deux mots. Néanmoins il y a entre les deux une grave différence de conditions. Dans *père* commençant une phrase (dans une interpellation par exemple), le mouvement de fermeture des lèvres est exécuté sans que rien précède, sans qu'aucun autre son soit interrompu par là. Au contraire, dans *repère*, le mouvement de rapprochement des lèvres succède aux mouvements articulatoires d'où résulte le *e* et les interrompt. Les mouvements sont sensiblement différents dans les deux cas. Le *p* intervocalique est autre chose que le *p* initial de mot devant voyelle. On conçoit donc que le traitement des consonnes intervocaliques

puisse différer de celui des consonnes initiales devant voyelle. Et c'est en effet ce que l'on observe d'une manière plus ou moins étendue dans un grand nombre de langues, sinon dans toutes.

Les langues celtiques fournissent des exemples clairs et instructifs, que M. Pedersen a bien mis en évidence dans ses recherches sur l'aspiration en irlandais. On fera ici abstraction du gaulois qui a de bonne heure cessé d'être parlé et sur lequel on n'a que des données partielles et insuffisantes, et où néanmoins il n'est pas impossible de mettre en évidence la tendance à l'altération des consonnes intervocaliques. Restent deux groupes : le groupe gaélique (en Irlande, en Écosse et dans quelques îles) et le groupe brittonique (dans le Pays de Galles, la Cornouaille anglaise et l'Armorique française). Dans les deux, les occlusives intervocaliques du celtique commun subissent une altération que ne connaissent pas les mêmes occlusives au commencement d'une phrase ou d'un groupe de mots autonome. Mais le résultat de l'altération n'est pas le même dans les deux groupes. En gaélique les occlusives intervocaliques sont devenues des spirantes, par exemple, *t* est devenu *th* (le *th* dur anglais), *d* est devenu *dh* (le *th* doux anglais), etc. ; ainsi c'est *athir* « père » qui, en vieil irlandais, répond au latin *pater*, *brāthir* qui répond au latin *frāter*, etc. En brittonique au contraire, les occlusives sourdes sont remplacées par les occlusives sonores correspondantes, ainsi *t* par *d* ; c'est donc *d* qui répond au *th* irlandais, et « frère » se dit en gallois *brawd* (pluriel *brodyr*), en cornique *broder*, en moyen breton *breuzr* (pluriel *breudeur*) ; les occlusives sonores deviennent spirantes, comme en gaélique ; par exemple à l'irlandais *bodar* « sourd » (dont le *d* note ici une spirante sonore) le gallois répond par *byddar*, le cornique par *bodhar*, le breton armoricain par *bouzar*. Ce contraste entre le brittonique et le gaélique n'a rien de fortuit ; on constate en effet que le brittonique maintient fidèlement des sonores anciennes que le gaélique assourdit, c'est-à-dire prive de leurs vibrations glottales : le *y* et le *w* celtiques communs sont conservés en brittonique, l'un sous la forme ancienne *y*, l'autre sous la forme *gw* ; en revanche, l'irlandais

ayant assourdi *y* et *w*, le premier a disparu entièrement, et le second s'est maintenu sous la forme de la spirante labio-dentale *f*; ainsi au latin *uērus*, le gallois répond par *gwir* et le breton armoricain par *guir*, tandis que l'irlandais a *fir*; le *y* initial du latin *iuuenis*, de l'allemand *jung*, etc. n'est représenté par rien dans l'irlandais *ōac* « jeune », mais est conservé dans gallois *ieuanc*, breton *iaouanc*.

Les occlusives ne sont naturellement pas les seules articulations qui aient un traitement particulier en position intervocalique. Toutes les consonnes sont exposées à des altérations analogues en cette position. On peut signaler telle langue où *dj* (le *j* notant ici et dans toute la suite *j* français) initial se maintient, mais où *dj* intervocalique devient *j*, ainsi la langue de l'Avesta (ancien iranien), où l'on a *djainti* « il frappe », mais *ajayō* « les serpents », représentant un plus ancien *adjayō*. Il y a de même telle langue où *s* initiale subsiste en principe, mais où *s* intervocalique passe à *h*; ainsi l'irlandais , où l'on a par exemple *sen* « vieux », comme on a latin *senex*, mais *aue* « de l'oreille » représentant *ausesos* ; l'intermédiaire a été *auhehòs*, d'où *aueos*, *auios*, *aue*.

La formule générale de l'altération des intervocaliques est malaisée à donner. Un très grand nombre des faits se laissent ramener à une diminution du degré de fermeture de la consonne ; c'est ce qui arrive quand *t* devenant spirant perd son occlusion, quand *dj* devenant *j* perd son très bref élément occlusif, quand *s* devenant *h* se réduit à un simple souffle. Cette tendance à l'ouverture des consonnes intervocaliques se conçoit aisément : les voyelles sont, par définition, des éléments ouverts ; on tend à réduire le plus possible le mouvement de fermeture entre deux phonèmes naturellement ouverts ; c'est un phénomène d'inertie. Quant à la sonorisation (passage de *t* à *d*, etc.), elle se conçoit également bien : d'une part, elle se rapproche du phénomène précédent en ceci que les sonores telles que *d* étant prononcées avec une occlusion moins ferme que les sourdes telles que *t*, le passage de *t* à *d* équivaut en quelque mesure à une diminution de fermeture ; d'autre part, les voyelles étant essentiellement sonores, c'est-à-dire pourvues de vibrations glottales, on tend à

maintenir la fermeture de la glotte d'où résultent les vibrations, même durant la prononciation de la consonne qui sépare les deux voyelles ; ici encore il s'agit d'un phénomène d'inertie.

Ce principe fait comprendre combien les langues peuvent différer au point de vue du traitement des intervocaliques. Dans une langue où les consonnes qui séparent les voyelles ont leur mouvement de fermeture bien marqué et suivi d'un silence appréciable, comme l'italien par exemple, le traitement des consonnes intervocaliques diffère peu de celui des initiales, et en effet, au moins après voyelle accentuée, les occlusives latines sont demeurées inchangées en italien ; au contraire, dans une langue où le mouvement de fermeture est peu appréciable, les consonnes intervocaliques tendent à subir un maximum de changements. Parmi les langues romanes, le français offre ainsi l'état inverse de celui de l'italien. Soit un latin *sapone(m)*, on a en italien *sapone*, avec le *p* conservé, et le même *p* se retrouve dans le roumain *sapun* ; le provençal *sabo* et le portugais *sabão* ont *b* ; le français *savon* et l'engadin *savun* ont un degré d'altération de plus, la spirante *v*.

Mais la tendance à l'altération des intervocaliques est néanmoins très générale. Ainsi l'italien garde l'occlusive sourde après l'accent, mais il la transforme en occlusive sonore après une voyelle inaccentuée, *fócum* y donne *fuoco*, mais *pacáre* y donne *pagar*. Parmi les anciennes langues indo-européennes, le grec, le slave et le lituanien se distinguent par une conservation fidèle des consonnes intervocaliques en tant que telles. Néanmoins, là où interviennent certaines circonstances favorables, la débilité des intervocaliques se manifeste. En grec, il existe une sorte de *w* qu'on nomme digamma et qui est en voie d'élimination dans la plupart des dialectes dès une époque préhistorique ; or, partout où l'on peut se rendre compte du développement des faits, on voit que la disparition du digamma intervocalique a précédé de beaucoup celle du digamma initial, ainsi à Delphes, en Crète, etc. Le slave doit à sa remarquable conservation des intervocaliques son aspect singulièrement archaïque : si le mot russe *nebesa* « les cieux » ne le cède pas en archaïsme au mot

sanskrit védique *nabhasā* c'est que le *b* et le *s* intervocaliques y sont intacts. Néanmoins le yod intervocalique est atteint en slave dès les plus anciens textes : *-ajego* tend à se contracter en *-aago*, *-ago*. La tendance à l'altération des intervocaliques est donc l'une des plus générales de toute la phonétique. Il n'entre pas dans le plan de cet article de rechercher les conditions physiologiques ou psychiques du fait; M. Jac. van Ginneken a du reste dit l'essentiel à cet égard dans ses *Principes de linguistique psychologique*, p. 448 et suiv. Mais les effets de cette tendance sont très inégaux suivant les langues.

Les exemples précédents, choisis parmi les plus clairs, montrent qu'il existe une phonétique générale. Les lois de la phonétique générale permettent d'interpréter les faits particuliers à chaque langue ; et il importe de les formuler explicitement, quand ce ne serait que pour éviter de faire appel inconsciemment à des principes de ce genre dont on n'aurait pas éprouvé la valeur par un examen approfondi. Beaucoup de linguistes, accoutumés par leur éducation de philologues aux précisions rigoureuses et aux règles presque toujours nécessaires de la grammaire descriptive et historique, ne voient pas sans inquiétude une discipline qui ne détermine que des possibilités et qui, ne pouvant jamais épuiser les faits de toutes les langues à tous les moments, doit procéder par induction en s'appuyant d'une part sur certains faits particulièrement nets et caractéristiques, de l'autre sur les conditions générales où ces faits se produisent. La linguistique générale est dans une large mesure une science a priori, comme l'a bien vu M. Séchehaye par exemple. Mais quelles que soient les difficultés de la linguistique générale ainsi comprise et quels qu'en soient les dangers, on n'a pas le droit de refuser de la faire. Car ne pas la faire explicitement, c'est se résigner à la faire implicitement, sans garanties et sans contrôle.

La linguistique générale ainsi comprise ne se suffit pas à elle-même. Elle repose sur la grammaire descriptive et historique à laquelle elle doit les faits qu'elle utilise. L'anatomie, la physiologie et la psychologie peuvent seules expliquer ses lois, on l'a vu par les exemples cités, et les considérations tirées de ces

sciences sont souvent utiles ou nécessaires pour donner une valeur probante à un bon nombre de ses lois. Enfin ce n'est que dans des conditions spéciales à un état social déterminé et en vertu de ces conditions que se réalise telle ou telle des possibilités déterminées par la linguistique générale. On voit ainsi quelle est la place de la linguistique générale, entre les grammaires descriptives et historiques d'une part, qui sont des sciences de faits particuliers, et l'anatomie, la physiologie, la psychologie et la sociologie, qui sont des sciences plus vastes dominant et expliquant entre autres choses les phénomènes du langage articulé.

CONVERGENCE

DES

DÉVELOPPEMENTS LINGUISTIQUES [1]

Quand une langue se différencie, comme l'ont fait par exemple à date historique le latin et l'arabe, les résultats de la différenciation varient à l'infini dans le détail matériel des faits, mais les lignes générales du développement sont la plupart du temps les mêmes. Si donc les ressemblances que présentent aujourd'hui les langues néo-latines ou les langues néo-arabes entre elles proviennent en partie de ce que ces langues ont conservé à un certain nombre d'égards l'état latin ou l'état arabe ancien, leurs ressemblances proviennent aussi en partie de ce qu'elles ont modifié dans un même sens l'état de choses de l'époque de communauté.

Quand on envisage le développement linguistique durant une période plus étendue, on observe des faits analogues.

L'indo-européen commun se parlait à une date relativement ancienne qui ne saurait être postérieure au début du second millénaire avant l'ère chrétienne. Entre l'indo-européen et les langues actuellement parlées qui sont des formes prises par cet idiome au cours du temps, comme le français, l'anglais, le persan, il y a des différences de détail infinies. Dans ces langues modernes, on ne reconnaît plus au premier abord l'unité d'origine, qui est certaine ; mais elles offrent des concordances frappantes dans leur structure générale.

L'arabe vulgaire, tel qu'on le parle aujourd'hui du Maroc à l'Égypte et à la Syrie, se distingue partout d'une même manière

1. *Revue philosophique*, t. LXXXV, nº de février 1918.

du sémitique commun, dont la comparaison du vieux babylonien, de l'arabe classique, de l'hébreu, etc. permet de restituer en une certaine mesure l'état ancien comme la comparaison du sanskrit, du grec ancien, du latin, etc., permet de restituer l'indo-européen. Et, chose remarquable, le passage du babylonien au néo-babylonien qu'on suit à travers la succession des textes est à plusieurs égards parallèle à celui qu'on observe de l'arabe classique aux parlers arabes modernes. Les langues sémitiques dont l'évolution a été relativement rapide comme l'hébréo-phénicien ou l'araméen, montrent, dès une date antérieure à l'époque chrétienne, un état d'altération de l'état sémitique commun analogue à l'altération qu'a subie l'arabe entre le VII^e^ siècle après J.-C. et l'époque moderne.

Voici un exemple de ces développements parallèles qui est à la fois simple et saisissant. Le trait caractéristique de la grammaire de l'indo-européen commun consiste en ceci que les relations soutenues par les mots dans la phrase sont exprimées par la forme même de ces mots ; chaque mot a autant de formes qu'il joue de rôles distincts dans la phrase. Par exemple l'indo-européen exprimait par une forme spéciale du nom du « père » le fait que ce nom sert de complément à un autre nom ; quand il dit *domus patris* ou *patris domus* — les deux ordres étant possibles — le latin ancien conserve exactement un état de choses indo-européen; or, ceci se dit en français *la maison du père*, en anglais *the house of the father*, en persan *mân i pidar*. Ces expressions se distinguent de l'expression indo-européenne par deux traits essentiels : d'une part, l'ordre des mots y est fixe, au lieu d'être libre, et, par là même, il est un moyen d'indiquer que le second nom est le complément du premier ; de l'autre, la forme du nom du « père » employée comme complément d'un nom est la même que celle qui est employée pour dire « le père est venu » ou « j'ai vu mon père » ou « je l'ai demandé à mon père », alors que, dans ces trois cas, le latin aurait employé trois autres formes : *pater*, *patrem*, *patri*. Mais, dans les trois langues en question, le rapport qui n'est pas exprimé par la forme du nom l'est par un mot accessoire, en même temps que par l'ordre

des mots : *de* en français, *of* en anglais, *i* en persan ; ces trois mots n'ont rien de commun entre eux ; le *de* français a une autre origine et un autre aspect que l'*of* anglais, et l'*i* persan a même une origine historique d'une espèce tout autre que *de* et *of*. Du reste l'anglais dit communément *the father's house*, où *s* se comporte comme un mot accessoire, tout en ayant une origine autre que *of* ; l'ordre des mots est autre que dans *the house of the father*, mais il est également fixe dans cette manière de s'exprimer. Ainsi en français, en anglais, en persan, on observe le même changement de structure par rapport à l'indo-européen : un ordre fixe des mots est substitué à un ordre libre, et un mot accessoire exprime la fonction ; mais si l'ordre est toujours fixe, il n'est pas le même dans toutes les langues et dans toutes les constructions ; et, s'il y a des mots accessoires, ces mots sont divers et d'origines diverses.

On peut multiplier les exemples de ce genre ; on rencontrera toujours le même type de faits : parallélisme des changements de structure générale, divergence des innovations portant sur les moyens matériels d'expression.

Dans les cas où une langue se différencie sans changer de type, le parallélisme des développements indépendants aboutit souvent à des résultats identiques, même dans le détail matériel de la forme. Ainsi la première personne du présent des verbes admettait en indo-européen deux finales suivant les cas ; ces deux finales sont bien conservées en grec ; il y a une finale *-ō* dans des verbes comme *legō* « je dis », et une finale *-mi* dans des verbes comme *didōmi* « je donne » ; c'est la distinction des verbes en *-ō* et des verbes en *-mi*, bien connue de ceux qui ont étudié quelques éléments du grec. La finale *-ō* avait plusieurs inconvénients ; elle manquait de corps, surtout dans les langues où les voyelles des syllabes finales tendent à s'altérer, ce qui arrive très souvent ; il est résulté de là une tendance à étendre l'emploi de *-mi* aux dépens de celui de *-ō*. Or, cette tendance s'observe dans une grande partie du domaine indo-européen : dans le groupe indo-iranien, toutes les premières personnes ont reçu *-mi* au présent, et, si un heureux hasard n'avait pas fait conserver quelques textes en un

vieux dialecte iranien où le représentant de l'ancien *-ō* a subsisté, on ne pourrait même pas démontrer que l'indo-iranien a connu autre chose que des premières personnes en *-mi* ; l'arménien a généralisé l'emploi de *-mi* (sous la forme *-m*) dans tous les présents ; l'irlandais a aussi étendu l'emploi de l'ancien *-mi* (tout en conservant le type en *-ō* en certaines conditions), et il a par exemple *scaraim* « je me sépare », tandis que, dans le même type de verbes, le latin a *amō* ; en germanique, le vieux haut allemand a la première personne *salbōm* « j'oins », tandis que le gotique a encore *salbo*, dans le même sens.

En slave, les développements parallèles de ce genre sont nombreux, parce que, la langue ayant conservé en général un type grammatical archaïque, il n'y a pas eu de changements généraux de structure qui entraînent des modes d'expression spéciaux à tel ou tel idiome. La première personne du présent était presque partout en *-o* nasal en slave commun, — cet *o* nasal résultant déjà d'une altération de l'ancien *-o* long final faite pour mieux marquer la forme — et l'ancien *-mi* n'était conservé que dans cinq ou six verbes slaves. Mais, l'un de ces verbes, le verbe signifiant « avoir », était très employé, et grâce à certaines circonstances dans l'histoire desquelles il est inutile d'entrer ici, l'ancienne finale *-mi*, réduite à *-m*, a été introduite d'abord dans les verbes en *-a-* où elle apparaît dès le moyen âge dans tous les dialectes slaves autres que le russe, puis, dans d'autres types verbaux, dans quelques dialectes seulement ; on suit dans la succession historique des textes le développement du fait, notamment par le tchèque et par le serbe, qui sont les langues où l'extension de *-m* a été le plus grande. Les langues slaves présentent ainsi un grand nombre d'innovations communes à plusieurs dialectes, parfois presque à tous les dialectes, et qui sont exactement les mêmes dans ces divers domaines ; quelques-unes de ces innovations, comme celle qui porte sur les formes employées avec les noms de nombre, sont inattendues au plus haut point, et pourtant elles se sont réalisées indépendamment dans plusieurs langues slaves longtemps après que ces langues avaient cessé de former une unité, et sans qu'il y ait eu influence de l'une de ces langues sur les autres.

On peut illustrer le principe par autant de faits particuliers qu'on le voudra. Tous ces faits se ramènent à une même formule :

Quand une langue se différencie en parlers distincts, celles des innovations réalisées dans chaque parler qui ne tiennent pas à des conditions propres à ce parler sont ou identiques ou du moins orientées en une même direction.

*
* *

De ce principe procède la régularité du développement linguistique, qu'on constate en fait.

Le langage est soumis à des conditions d'existence qui en déterminent le développement. Ces conditions expliquent le parallélisme des innovations qu'on vient de noter.

Chacune des notions exprimées par le langage l'est au moyen d'un mot auquel est associé un sens défini. Toutes les langues procèdent par mots.

Mais les choses à exprimer sont trop multiples, trop variées et trop nuancées pour qu'il puisse y avoir autant de mots que de notions. On ne réussit donc à tout dire qu'en faisant subir aux mots des variations ou des groupements destinés à rendre les nuances.

Toutes proportions gardées, on peut dire que le langage procède avec les mots comme le typographe procède avec les caractères mobiles qui lui servent à composer des mots infiniment divers. Les *mots* d'une langue se laissent combiner avec d'autres mots de toutes sortes de manières : les combinaisons sont soumises aux règles de la grammaire ; mais les règles n'empêchent pas les combinaisons possibles d'être innombrables.

D'autre part, on ne parle que pour établir un lien entre deux notions, pour dire par exemple que *Pierre est bon* ou que *Pierre vient*. Si le sujet auquel on attribue quelque qualité ou quelque action est connu, on peut ne pas l'indiquer expressément ; si l'on sait de qui il s'agit, le latin peut dire : *bonus est* ou *venit*. Mais, le plus souvent, on est amené à exprimer dans la phrase à la fois le « sujet » et le « prédicat ».

La grammaire d'une langue est l'ensemble des procédés par lesquels, dans cette langue, on groupe les mots.

Il y a des langues où la forme des mots est invariable ; la comparaison avec les caractères mobiles du typographe est alors assez satisfaisante. Il y a d'autres langues — et l'indo-européen en est le type le plus extrême — où les mots affectent des formes diverses suivant le rôle qu'ils jouent dans la phrase : on peut comparer les langues de cette sorte à un manuscrit où chaque lettre admet des formes diverses suivant les lettres auxquelles elle se joint dans l'écriture.

L'expérience montre que le nombre des procédés possibles est petit et que les catégories grammaticales sont, au moins dans leurs traits généraux, peu nombreuses.

Soit, par exemple, le verbe. Tout langage exige qu'on exprime de quelque manière les différences entre les personnes qui parlent, à qui l'on parle ou dont on parle. Il faut qu'on puisse distinguer entre : *je dis, tu dis, il dit, nous disons, vous dites, ils disent.*

Le détail des distinctions peut varier. Ainsi la plupart des langues de l'Europe tendent à confondre *tu dis* et *vous dites* en une expression unique, par suite d'une habitude de politesse : en français la distinction entre *tu dis* et *vous dites* ne subsiste que dans le langage familier ; en anglais elle est abolie.

Certaines langues de demi-civilisés distinguent deux manières de dire : « nous disons », suivant qu'il s'agit de « moi et toi (ou vous) » (pluriel inclusif) et de « moi et lui (ou eux) » (pluriel exclusif). D'autres langues distinguent entre *nous* (*deux*) *disons*, *vous* (*deux*) *dites*, *eux* (*deux*) *disent* et *nous* (*plus de deux*) *disons*, *vous* (*plus de deux*) *dites*, *eux* (*plus de deux*) *disent* : cette distinction du duel et du pluriel disparaît partout avec le progrès de la civilisation, comme celle des formes inclusives et exclusives. Partout le progrès de la civilisation tend à détruire les formes demi-concrètes et à ne laisser subsister que l'opposition abstraite de deux catégories, celle de l'unité et celle de la pluralité.

A quelques détails près, dont une tendance générale commande du reste le développement comme on le voit, les catégories rela-

tives au nombre et à la personne dans les verbes sont partout les mêmes, et cette identité résulte des conditions d'existence du langage.

Quant aux procédés employés pour exprimer ces catégories, il n'y en a que deux. L'un, qui est celui de l'indo-européen commun conservé dans les anciennes langues du groupe telles que le sanskrit, le grec ancien, le latin, etc., consiste à faire varier le verbe et à lui donner autant de formes qu'il y a de catégories à exprimer : ainsi en latin *dico, dicis, dicit, dicimus, dicitis, dicunt*. L'autre consiste à employer des mots accessoires, qui indiquent chaque catégorie ; c'est ce qu'on fait en anglais quand on dit : *I say, we say, you say, they say*. Enfin on peut employer à la fois les deux procédés ; ainsi le français ne distingue pas (sauf dans l'orthographe) les trois personnes du singulier du verbe, autrement que par le mot accessoire qui précède : *je dis, tu dis, il dit* (et ainsi dans tous les verbes français sauf *j'ai* et *je suis* où *ai* et *suis* se distinguent de l'*a* de *tu as, il a* et de l'*e* de *tu es, il est*) ; mais les trois personnes du pluriel ont chacune une forme propre : *nous disons, vous dites, ils disent*, et ainsi toujours, à ceci près que la troisième personne du pluriel ne se distingue souvent pas de celle du singulier autrement que par l'orthographe : *il aimait* et *ils aimaient, il aime* et *ils aiment* ne diffèrent que par le mot accessoire précédent.

Or, chacun des deux procédés tend à se détruire par l'effet de l'usage.

Le procédé indo-européen commun, conservé dans le type latin *dico, dicis*, etc., a le défaut grave de dissimuler l'unité du mot. Il y a en anglais un mot *say* signifiant « dire » ; l'addition de *-s*, dans *he says* « il dit » n'en dissimule guère la parfaite unité. En latin, il n'y a pas de mot signifiant « dire » ; le présent du verbe ayant ce sens est l'ensemble de *dico, dicis*, etc., sans qu'une des formes ait plus que les autres un titre à dominer. Quand on veut énoncer un verbe latin, on le fait soit au moyen de l'infinitif présent, *dicere*, soit au moyen de la première personne du singulier du présent, *dico* ; mais ce n'est qu'affaire de convention. En sanskrit, où la structure linguistique est la même à ce point

de vue, on emploie une convention différente ; et l'on y cite les verbes à la troisième personne du singulier, ce qui est aussi arbitraire, mais pas plus.

Contre cette insaisissabilité du mot, toutes les langues du groupe indo-européen n'ont cessé de réagir. Ce serait à peine une exagération trop forte de dire que l'histoire des langues indo-européennes se résume essentiellement en un effort pour passer du mot-forme, existant seulement à l'état de formes fléchies multiples dont chacune a une valeur particulière, au mot existant isolément et toujours semblable à lui-même. L'état de choses indo-européen est aussi loin que possible du mot isolable ; une langue comme l'anglais y est arrivée presque entièrement, et les cas où un mot affecte plusieurs formes suivant le sens à exprimer y sont devenus exceptionnels : on compte en anglais les cas d'opposition comme celle entre le singulier *foot* « pied » et le pluriel *feet* « pieds ».

La puissance de la tradition est grande. Depuis quelque quatre mille ans qu'elle agit dans les langues indo-européennes, la tendance au changement n'a pleinement abouti nulle part, et certaines langues, les langues slaves notamment, n'ont aujourd'hui encore que des mots fléchis suivant le vieux type indo-européen. Mais l'action de la tendance a été universelle et constante. Ses effets se voient partout. Et c'est l'action constante de cette tendance qui rend compte du fait que les langues indo-européennes se sont développées parallèlement les unes aux autres.

La flexion a un autre défaut que celui de masquer l'unité du mot : c'est de n'être pas expressive. Le mot ne ressort guère dans le latin *dico* ; mais l'indication de la personne et du nombre n'y ressort pas non plus bien fortement. Si l'on veut insister sur la personne, il faut l'indiquer par un mot séparé, et l'on dit en latin : *ego dico*, qui ne signifie pas « je dis », mais « moi, je dis » ou « c'est moi qui dis ». Toutes les anciennes langues du groupe indo-européen procèdent ainsi.

Enfin la flexion avait un troisième défaut : elle variait suivant les mots auxquels elle s'appliquait ; en latin, *dico*, *dicis*, ou *amo*, *amas*, ou *moneo*, *mones*, ou *capio*, *capis*, ou *sum*, *es* ne se com-

portent pas de la même manière. En revanche, en employant des petits mots accessoires pour marquer la personne, on établit l'unité dans le moyen de l'exprimer. Le français caractérise essentiellement la personne par un mot accessoire, toujours le même, ainsi dans : *je dis, tu dis* ; *j'aime, tu aimes* ; *j'avertis, tu avertis* ; *je prends, tu prends* ; et une opposition comme celle de *je suis, tu es* — où d'ailleurs figurent *je* et *tu* — y est une survivance exceptionnelle.

D'une manière générale, les langues indo-européennes sont amenées à indiquer par des mots autonomes les notions que l'ancienne grammaire indo-européenne rendait d'une manière implexe par les formes de la flexion. A la tendance, toute intellectuelle, vers l'unité du mot et vers l'unité d'expression de la catégorie grammaticale se joignait, agissant dans le même sens, une tendance, de caractère émotionnel, vers une expression plus intense.

Mais l'usage diminue vite l'efficacité d'un tour expressif. Quand on disait en latin *ego dico*, c'était pour insister sur l'idée de « moi » ; mais pour peu que ce tour ait été répété, la valeur d'insistance de *ego* a baissé. Et, comme on parle d'ordinaire pour agir et qu'on s'efforce en général d'attirer l'attention de l'interlocuteur, on a été amené à employer de plus en plus ce tour qui devait perdre par là même sa puissance initiale. Dès lors, puisque *ego* indiquait la personne, la flexion devenait de moins en moins utile.

Dans le passage du latin au français, les pronoms qui, en latin ancien, étaient des mots autonomes, sont devenus de simples indices de la personne. Ils ont même perdu toute autonomie : *je, tu, il, ils* ne sont plus des mots énonçables séparément, ayant un rôle par eux-mêmes ; *je dis, tu dis, il dit,* où il y a deux éléments séparables, ne sont pas plus expressifs aujourd'hui que *dico, dicis, dicit* ne l'étaient en latin. Et les variations de flexion de *nous disons, vous dites, ils disent* ne sont plus que des survivances traditionnelles ; elles subsistent parce qu'elles se transmettent et qu'aucun accident ne les a encore détruites ; mais il n'y a pas plus de raison de distinguer la forme du verbe dans *nous disons, vous dites,* que dans *je dis, tu dis,* et dans *je joue, nous*

jouons que dans *il joue, ils jouent*, où les différences ne sont que graphiques.

D'ailleurs la confusion des deux formes de *il*, pour le singulier et pour le pluriel, est un grave défaut de la langue française. Si le français n'était pas fixé par une tradition rigide, la confusion de *il joue* et *ils jouent* s'éliminerait, soit par une différenciation de la forme du pronom, qui répondrait à la tendance générale du développement des langues indo-européennes, soit par une différenciation de la forme du verbe, qui répondrait à l'état de choses actuel du français; le français a en effet conservé une flexion verbale au pluriel; et sur le modèle de *ils sont, ils font*, etc., beaucoup de parlers populaires ont actuellement *i(ls) jouont* au pluriel. On aperçoit ici comment les tendances générales du développement linguistique sont croisées par des tendances propres à chaque langue en particulier, et, comment, par suite, ces tendances générales n'aboutissent que lentement et après avoir été arrêtées longtemps par des innovations accidentelles, et aussi comment il se fait qu'elles n'aboutissent pas partout à la même date ni sous la même forme.

La tendance à remplacer la flexion par des mots accessoires a été favorisée dans les langues indo-européennes par un détail de structure de la grammaire et par un principe de phonétique générale. La flexion indo-européenne procédait essentiellement par des éléments ajoutés ou soudés à la fin des mots, par suffixes et par désinences, pour employer les termes techniques. C'est donc la fin du mot qui renfermait les caractéristiques de flexion. Or, l'observation montre que, dans un mot de plus d'une syllabe, la syllabe finale tend à s'affaiblir dans la prononciation, et finalement à disparaître. Le français, où la syllabe qui est d'ordinaire prononcée un peu plus forte, plus longue et plus haute que les autres est la syllabe finale, ne laisse pas voir au premier abord les effets de cette tendance. Mais, s'il est parvenu à cet état, c'est précisément parce que la syllabe finale des mots latins y a été entièrement détruite au cours du temps; là où le latin avait *focum*, et où l'italien a encore *fuoco* et l'espagnol *fuego*, le français a *feu*, sans trace de ce que renfermait la syllabe finale du mot latin. Si,

dans un mot comme *raison*, l'accent est sur la fin du mot, c'est parce que le *-e(m)* final du latin *rationem* n'a pas subsisté en français ; l'italien qui est resté plus fidèle à l'état ancien, a l'accent sur la syllabe pénultième dans *raziône* . L'affaiblissement progressif des finales et la destruction qui a été la conséquence dernière de cet affaiblissement ont beaucoup contribué à développer l'usage de mots accessoires : moins la finale était nette, moins la flexion était claire, et plus l'emploi de mots accessoires était utile. Une langue comme l'anglais, qui a perdu toutes les anciennes finales, ne peut plus exprimer les relations grammaticales que par l'ordre des mots et par des mots accessoires.

Le développement ne conduit pas toujours à substituer des mots isolés non fléchis, déterminés par des mots accessoires, à des mots fléchis. Le développement inverse se fait en même temps. En effet, les mots accessoires perdent de plus en plus leur sens propre et leur autonomie, et ils tendent à faire partie intégrante des mots qu'ils déterminent. Par exemple, en français, *je, tu, il* dans *je dis, tu dis, il dit* n'ont plus le caractère de mots ; ces soi-disant pronoms n'existent qu'auprès des verbes, et, si on les isole, ils n'ont aucun sens ; aucun Français ne pense *je, tu, il* séparément ; et, si l'on a à exprimer la personne d'une manière spéciale, on recourt à *moi, toi, lui* : *moi je dis, toi tu dis, lui il dit*. Au fur et à mesure qu'elle se détruit, la flexion tend à se reconstruire. L'histoire du futur français fournit un bon exemple : les langues romanes ont perdu le vieux futur *dicam, dices*, etc. ; on l'a remplacé, en France notamment, par un groupe *dicere habeo* « j'ai à dire », qui comprend deux mots ; le représentant du mot accessoire *habeo*, réduit à *ai*, s'est fondu avec le verbe, si bien que le français a de nouveau une flexion *je dirai, tu diras*, où personne ne reconnaît le verbe *avoir*.

En ce qui concerne les cas où le développement indépendant de parlers séparés aboutit spontanément à des résultats identiques, on ne pourrait montrer comment ces résultats sont nécessaires qu'en entrant dans des détails echniques. Mais il serait aisé d'établir que l'identité des résultats provient de l'identité des conditions particulières dans lesquelles a lieu le développement.

*
* *

On a souvent dit que les innovations linguistiques sont des créations individuelles généralisées. Les théoriciens qui insistent sur ce fait le font pour mettre en évidence la part d'invention individuelle et de libre choix qu'il y a dans le développement du langage. Et ils signalent avec force le grand rôle de l'imitation dans le changement linguistique.

Pour autant qu'il s'agit de vocabulaire et de tours de phrase, la part d'invention individuelle n'est pas niable. Mais ce sont les parties du langage où il ne se produit pas de faits de convergence. En dehors de quelques accidents fortuits qui ne peuvent pas ne pas se produire comme l'identité du mot signifiant « mauvais » en anglais et en persan, *bad*, ou la ressemblance de *feu* en français et de *feuer* en allemand, sans que ces mots aient rien de commun, deux mots qui ont le même sens ne se recouvrent d'une langue à l'autre que s'ils ont une origine commune ou si l'une des langues a emprunté à l'autre. Le développement indépendant de deux langues ayant même origine les amène à avoir des vocabulaires de plus en plus divergents. Par exemple le vocabulaire français et le vocabulaire persan diffèrent presque du tout au tout.

Dans les parties systématiques du langage, où la convergence s'observe fréquemment, à savoir la phonétique et surtout la grammaire, l'invention individuelle ne saurait guère intervenir, comme on peut le prévoir aisément si l'on observe le fait, familier à tout le monde, que l'on articule les sons et que l'on fait les formes grammaticales sans avoir conscience des procédés employés ; c'est ce que montrent immédiatement les faits qu'on vient d'analyser.

Sans doute, une langue n'existe pas en dehors des gens qui la parlent; et il n'y a d'innovations que celles qui ont été faites par des individus qui se servent de la langue. En ce sens, il est licite d'affirmer l'origine individuelle des innovations; mais il est inutile de formuler une vérité aussi évidente, un truisme aussi naïf.

Ce qui est essentiel, ce sont les forces qui déterminent les changements. Or, ces forces agissent sur la collectivité des gens qui parlent une même langue. Et c'est parce qu'ils se trouvent dans les mêmes conditions et qu'ils subissent les mêmes actions, que les sujets parlants admettent les mêmes innovations. Ces actions sont en partie universelles, en partie spéciales à certaines langues. Ce sont les actions universelles qui produisent les convergences : elles sont nombreuses et puissantes. Il y a aussi des actions particulières à telle ou telle langue, et ce sont ces actions qui déterminent les divergences de langues anciennement unes. Générales ou spéciales, ces actions n'aboutissent à un résultat que si elles s'exercent sur toute la collectivité, ou, du moins, sur une très grande partie de la collectivité.

La question de savoir en quelle mesure les actions collectives se manifestent spontanément chez les divers individus qui parlent une langue donnée n'est pas résolue. On n'a jamais réussi à observer un changement grammatical en voie de réalisation depuis le moment où il apparaît jusqu'au moment où la nouvelle forme est de règle. Et si, par fortune, on réussissait à faire une observation complète, ce ne serait jamais qu'un cas particulier dont on ne saurait tirer de conclusions générales. Toutefois, puisque l'on constate que des changements identiques, ou du moins très semblables, se réalisent indépendamment dans des dialectes semblables entre eux, mais entièrement séparés, comme l'étaient au moyen âge les dialectes slaves, on a par là la preuve que, si les mêmes conditions se rencontrent, les langues se développent non seulement en un même sens, mais aussi de la même manière jusque dans le détail matériel des formes. Les mêmes innovations se produisent donc indépendamment chez des individus différents pourvu qu'ils soient placés dans les mêmes conditions. Ceci ne prouve naturellement pas que les changements naissent spontanément chez chacun des sujets, et qu'il n'y ait pas, en une plus ou moins large mesure, imitation d'un sujet par un autre ; mais il en résulte au moins qu'ils sont susceptibles de naître indépendamment chez plusieurs sujets, et souvent même chez beaucoup de sujets. Et, quant aux sujets chez

lesquels les innovations ne naissent pas spontanément, ils sont out prêts à les accepter, parce qu'elles répondent à un besoin senti par eux. L'existence d'une tendance collective est ce qui domine tout ; peu importe le rôle que joue dans la réalisation des changements l'imitation. Grand ou petit, ce rôle n'est en tout cas qu'un élément accessoire, tandis que la tendance générale est le principe d'où tout dépend. L'historien, qui se plaît à suivre des faits particuliers, peut désirer connaître les procès de détail par lesquels se font les innovations grammaticales ; mais le linguiste qui a affaire avec le fait collectif du angage, se résigne aisément à les ignorer.

Par cela même que l'on constate des développements linguistiques semblables là où se rencontrent des conditions semblables, et des développements identiques là où se rencontrent des conditions identiques, on voit que les résultats dépendent de ces conditions, et qu'ils sont indépendants des procédés par lesquels ils se réalisent.

Les considérations qui précèdent ne visent pas à contester le rôle de l'imitation dans le langage. Il arrive souvent que des groupes sociaux, parfois des groupes très étendus, changent de langue, ou changent de parler. Après la conquête romaine, l'aristocratie gauloise a très vite accepté le latin, sans lutte linguistique. Les parlers locaux disparaissent aujourd'hui en France devant le français commun. Et une langue peut s'étendre, comme on le voit, à l'époque historique, par l'exemple du latin ou de l'arabe. Des manières de parler peuvent se généraliser : le parler des centres principaux de civilisation tend à remplacer dans le monde moderne les parlers locaux qui ont moins de prestige. Il y a là des faits qui sont de grande importance pour le développement linguistique, et qui procèdent uniquement de l'imitation. Mais ce n'est pas de ces faits — capitaux — qu'il a été question ici.

Il faut bien distinguer deux ordres de changements. Il y a, d'une part, les changements de langue et de parler qui consistent dans l'adoption d'une langue ou d'une manière de parler ayant un prestige. Et il y a, d'autre part, le changement linguistique, à l'intérieur d'une langue, changement qui provient de l'action de

tendances, les unes universelles, les autres spéciales à tel ou tel groupe de langues, à telle ou telle langue. C'est dans ce second cas, celui du changement linguistique, qu'ont lieu les faits de convergence étudiés ici.

Les convergences observées permettent d'établir que, en matière de changement linguistique, les innovations sont générales plutôt que généralisées, et que l'identité ou la parité des conditions où se trouvent les sujets parlants est le fait essentiel, l'imitation une chose secondaire.

LE PROBLÈME
DE LA
PARENTÉ DES LANGUES[1]

La dernière liste de langues publiée, celle qu'a donnée le regretté F. N. Finck, *Die Sprachstämme des Erdkreises*, en renferme plus de deux mille. Ce chiffre n'a, par lui-même, aucune importance : la liste n'a pas été faite suivant un principe un, et il ne serait pas possible d'en dresser une avec rigueur. En effet on peut convenir de compter autant de langues distinctes qu'il existe de parlers inintelligibles pour les sujets employant tout autre parler. Mais cette limite d'intelligibilité est fuyante : le provençal et le français sont assurément deux langues distinctes ; car pour qui ne sait que l'une de ces deux langues, l'autre est inintelligible à l'audition ; mais le provençal et le catalan sont-ils deux langues distinctes ? Au premier abord, les sujets qui parlent l'une de ces deux langues ne saisissent pas l'autre ; mais il faut peu d'efforts à un Provençal pour comprendre un Catalan et inversement. Il est donc impossible de dire combien il existe actuellement de langues. Une chose du moins est sûre : les langues distinctes se comptent par centaines, et il se parle actuellement beaucoup plus de mille idiomes assez différents les uns des autres pour rendre nécessaire l'usage d'interprètes.

Pour mettre un ordre dans cette variété, on a tenté de classer les langues d'après les traits généraux de leur structure grammaticale en langues isolantes, agglutinantes, incorporantes et flexionnelles. Mais on n'a pas réussi à poursuivre ce classement jusqu'au bout, et, dans la mesure où il a été fait, il s'est trouvé dénué de toute utilité soit pratique, soit scientifique ; c'est une

1. *Scientia* (*Rivista di scienza*), vol. XV, nº XXXV-3 (1914).

amusette dont aucun linguiste n'a pu tirer parti. Dans son petit livre sur les principaux types de langues, F. N. Finck n'a pu que décrire huit langues particulières arbitrairement choisies.

Fr. Müller, dans son grand exposé des principaux types de langues, et F. N. Finck, dans le petit ouvrage cité ci-dessus, ont été réduits à ranger les langues d'après les races d'hommes qui les parlent. Mais comme il n'y a aucun lien nécessaire entre la langue et la race et que personne ne soutient sérieusement qu'un certain type linguistique soit lié à un certain type somatique, ce procédé n'a été adopté que comme un pis-aller, et un pis-aller fâcheux. Car il éveille une idée fausse.

Toutefois, si l'on ne connaît aucun principe qui permette actuellement de classer toutes les langues, on a déterminé l'existence de plusieurs groupes linguistiques nets; personne ne doute qu'il y ait un groupe roman, composé de l'italien, de l'espagnol, du français, du roumain, etc. ; un groupe germanique, composé de l'allemand, de l'anglais, du danois, du suédois, etc. ; un groupe slave, composé du tchèque, du polonais, du russe, du serbe, du bulgare, etc. ; et ainsi de beaucoup d'autres. Les ressemblances entre les langues qui constituent chacun de ces groupes sont évidentes ; la valeur de ce classement se traduit pratiquement par le fait que si l'on possède l'une des langues du groupe, on acquiert plus facilement les autres. Il y a donc un classement linguistique qui s'applique d'une manière satisfaisante en certains cas. Dès lors il vaut la peine d'examiner si, en appliquant le procédé, on n'arriverait pas à classer toutes les langues.

Le principe de ce classement est simple. Toutes les langues actuellement parlées du groupe roman sont des transformations diverses d'une seule et même langue qui ne se parle plus, mais dont on a des monuments écrits, le latin. Les langues du groupe germanique sont des transformations d'une même langue commune, dont on n'a pas de monuments, mais dont on doit supposer l'existence ; les langues slaves sont les formes diverses prises par une seule et même langue slave commune non attestée. Plus on remonte dans le passé de chacune des langues germaniques, plus se manifeste la ressemblance ; et de même pour

le groupe slave ; si les documents conservés permettaient de suivre l'histoire plus loin en arrière, on finirait — et assez vite — par trouver l'identité complète, dans la mesure très approximative où une langue est jamais une. Les groupes de langues utilement constatables sont donc définis par l'unité d'origine. Dire que plusieurs langues appartiennent à un même groupe, c'est dire qu'elles sont des différenciations d'une langue plus ancienne : des langues *parentes* sont en réalité *une seule et même langue modifiée de manières diverses* au cours du temps.

Il y a donc une classification linguistique qui, en certains cas au moins, est satisfaisante, la classification « généalogique » ; elle a abouti pour beaucoup de langues à des résultats précis que Fr. Müller et F. N. Finck ont insérés au milieu de leur classification ethnique générale ; elle a permis, depuis une centaine d'années, de faire l'histoire des langues indo-européennes, des langues sémitiques, des langues finno-ougriennes, des langues bantoues, des langues indonésiennes, et sa fécondité se montre chaque jour plus grande ; il était permis de croire qu'elle ne donnerait plus lieu à des discussions et qu'on s'efforcerait seulement d'en poursuivre l'application. Mais elle a été dans les dernières années l'objet de discussions de principe. Il y a en effet des obscurités dont les unes proviennent de la manière dont les langues se transforment, et les autres tiennent à ce qu'il est malaisé de mettre en évidence des communautés d'origine. D'autre part, la difficulté de fait qu'on éprouve à faire entrer toutes les langues dans la classification généalogique a conduit certains linguistes éminents à ôter au principe de cette classification sa précision et sa rigueur ou à l'appliquer d'une manière inexacte. Il ne sera donc pas inutile de remonter à ce principe même et de montrer en quelle mesure une classification généalogique des langues est actuellement possible, jusqu'où elle peut être utile, et ce que l'on en peut espérer.

I

La formule suivant laquelle certaines langues sont les trans-

formations d'une langue plus ancienne recouvre des réalités historiques diverses et complexes. En effet les langues ne sont pas des objets ayant une existence matérielle autonome et se développant par eux-mêmes. Une langue est une institution propre à une collectivité sociale, et les modifications qu'elle subit sont liées à l'histoire de cette collectivité. Elle ne se transforme pas comme un manuscrit ou comme un outil. Elle résulte d'une activité mentale des sujets parlants. Elle n'est pas la même chez deux sujets parlant une même langue : il n'y a pas deux Français qui parlent exactement le même français. Sans doute il existe pour chaque langue une norme à laquelle les sujets parlants tendent à se conformer; mais cette norme est plus ou moins définie suivant les cas, et le sentiment qu'on en a est plus ou moins délicat. Des variations, même étendues, n'empêchent pas les gens de se comprendre entre eux. Par suite, il existe des communautés linguistiques où l'unité est très imparfaite ; et il n'y en a aucune où tous les sujets parlent d'une manière sensiblement identique. La formule : une langue est une forme modifiée d'une autre, n'a donc pas un sens linguistique précis, puisque ni l'une ni l'autre langues ne se laissent exactement saisir à aucun moment.

Du reste les langues ne se maintiennent qu'en se transmettant d'un individu à un autre. Or, les enfants qui apprennent à parler ne reçoivent pas la langue toute faite. Ils ne peuvent reproduire que ce qu'ils entendent, et il est inévitable que des nuances délicates échappent à leur attention. En constituant leur système linguistique avec ce qu'ils ont entendu et remarqué, ils ne reproduisent pas exactement le système linguistique des générations antérieures. Même quand ils emploient les mêmes formes, ils ne les sentent pas toujours de même. Par sa discontinuité naturelle, la transmission du langage donne lieu à des changements. Les enfants sont d'autant plus libres vis-à-vis des générations antérieures que la norme est moins précisément établie et que l'unité est moins réalisée dans la communauté où ils apprennent à parler.

La langue ne se transmet pas seulement de génération en gé-

nération. Sans cesse il se produit des situations qui engagent ou qui obligent certains sujets à employer, de préférence à leur parler maternel, une langue susceptible de servir à des communications avec un plus grand nombre d'hommes; sans cesse, il y a par suite des groupes d'hommes qui changent de langue. Mais on ne sait jamais parfaitement une langue étrangère ; on n'en saisit pas toutes les finesses; on n'en reproduit pas exactement tous les détails ; on n'en applique pas correctement tout le système. Chaque sujet qui emploie une langue étrangère apporte donc un trouble dans la communauté linguistique. Partout où il y a des populations d'origines diverses, la situation linguistique devient incertaine, et la liberté des enfants qui apprennent à parler s'accroît. Quand, ainsi qu'il arrive souvent, une population entière change de langue, il est inévitable que la langue nouvellement apprise, d'une part, ne soit pas en tous points conforme au modèle imité, et, d'autre part, soit pendant un certain temps moins stable que n'est une langue fixée par une longue tradition.

Quels que soient les procès, en partie mal éclaircis, par lesquels se réalisent les changements, une chose est sûre : entre l'aspect d'une langue à un certain moment et l'aspect de cette même langue à une époque postérieure, il y a souvent une différence profonde ; en quelques siècles, une langue peut devenir méconnaissable. Le français du XI^e^ siècle est du latin, à savoir le latin vulgaire qu'ont adopté vers le I^er^ siècle après J.-C. les populations de la Gaule du fait de la conquête romaine ; mais un Romain du I^er^ siècle n'aurait pas compris ce latin nouveau pas plus qu'un Français du XI^e^ siècle n'était capable d'entendre le latin ancien. Depuis, le français a continué de changer rapidement en s'éloignant de plus en plus du latin. Mais, si fort que la divergence grandisse, le français ne cessera pas pour cela d'être du latin transformé, aussi longtemps qu'il se parlera.

En effet, la définition de l'identité linguistique ne peut être que sociale : quelles que soient les différences de fait entre les sujets parlants, il y a langue une là où des individus, se comprenant entre eux, ont, d'une façon consciente ou inconsciente, le sentiment et la volonté d'appartenir à une même communauté

linguistique. La langue change au cours du temps ; mais, si l'on considère des sujets contemporains les uns des autres, il y a en un sens unité à chacun des moments successifs qu'on peut envisager. Entre la conquête de la Gaule par les Romains et l'époque actuelle, il n'y a eu aucun moment où les sujets parlants aient eu la volonté de parler une langue autre que le latin. Les habitants de la Gaule ont voulu acquérir le latin, et ils y sont parvenus ; le latin vulgaire dont le français est une transformation diffère très peu de celui qui a abouti aux autres langues romanes. Depuis ce temps, les habitants de la France actuelle n'ont jamais eu l'intention de parler une langue nouvelle ; ils se sont toujours efforcés de parler leur langue traditionnelle, et ils ont toujours eu le sentiment de parler une même langue. Une différence tranchée n'apparaît que si l'on envisage deux moments éloignés du développement. — Par le fait de la diversité des conditions historiques, une même langue peut subir des changements divers dans les différentes régions où elle est employée ; dès lors, malgré la persistance de la continuité avec le passé en chaque région, cette langue se brise en un certain nombre de langues qui progressivement deviennent distinctes. Ainsi le latin est devenu suivant les lieux l'italien, le provençal, le français, l'espagnol, le roumain, etc. — On voit par là que ce qui définit une parenté linguistique, c'est seulement un fait historique : une langue sera dite issue d'une autre si, à tous les moments compris entre celui où se parlait la première et celui où se parle la seconde, les sujets parlants ont eu le sentiment et la volonté de parler une même langue, soit que cette langue se soit transmise normalement de génération en génération, soit que certains groupes d'hommes l'aient adoptée à la place de leur ancien parler. Sont parentes entre elles toutes les langues issues ainsi d'une même langue. Ainsi la parenté de langues résulte *uniquement* de la continuité du sentiment de l'unité linguistique.

Le sentiment et la volonté qu'ont les sujets parlants d'employer une certaine langue sont chose inconsciente là où il n'existe qu'une langue : dans un milieu linguistique homogène, les sujets parlent comme on parle autour d'eux. Mais, dans les milieux où

il existe plusieurs manières de parler, on sait toujours qu'on se sert de tel ou tel système : un Breton qui pratique le français et le breton sait en chaque cas s'il parle français ou breton, et il ne mélange pas l'un des systèmes avec l'autre ; car les systèmes grammaticaux de deux langues sont, on le verra, impénétrables l'un à l'autre.

On peut être tenté de dire que la langue issue d'une autre plus ancienne, si elle n'est ni identique ni même semblable à celle-ci, est constituée du moins avec les mêmes éléments linguistiques. Mais il n'y a rien à tirer de là.

Tout d'abord ce fait n'a par lui-même guère d'intérêt. Sans doute le *-ai* du français *je ferai* et le *-ais* du français *je ferais* sont des formes modifiées du latin *habeo* et *habebam*. Mais ces mots latins sont si profondément altérés qu'on les reconnaît à peine ; même si *je ferai* était clair, le simple examen du français ne donnerait pas le moyen d'analyser *je ferais*. Et, s'il est vrai que *-ai* et *-ais* sont des éléments latins, ils ont pris un emploi qui n'a pas son équivalent en latin et que rien en latin ne laissait prévoir : entre les formes grammaticales françaises *je ferai* et *je ferais* et les mots latins *habeo*, *habebam*, il n'y a de commun qu'un fait matériel : *-ai* et *-ais* sont des déformations de *habeo* et de *habebam*. En tant que formes grammaticales, ce sont de véritables créations. Dans des conditions favorables pour le linguiste, comme celles offertes par les langues romanes, de pareilles créations se laissent expliquer. Mais, dans la plupart des cas, elles demeurent mystérieuses. La grammaire comparée explique bien les survivances d'un état de choses ancien ; mais elle échoue souvent quand il s'agit d'expliquer les innovations qui sont de véritables créations. Donc, si l'on considère deux moments du développement d'une même langue séparés par un large intervalle de temps ou par une période de transformations multiples et rapides, on n'arrivera peut-être à presque plus rien expliquer dans le second moment par la considération du premier moment. Les éléments du premier moment considéré seront méconnaissables dans le second.

En second lieu, les langues sont sujettes à emprunter des élé-

ments à des langues étrangères. Il y a en français moderne des mots nouvellement empruntés à l'anglais, à l'allemand, à l'italien, à l'espagnol, etc., et dont les sujets parlants savent bien qu'ils n'appartiennent pas au vieux fonds français. En ce sens toute langue est mélangée d'éléments étrangers. La part de ces éléments étrangers peut être grande ; le persan actuel a un vocabulaire dont, surtout dans la langue littéraire, la plus grande partie se compose de mots arabes ; le turc osmanli est plein de mots arabes et de mots persans ; l'anglais a pris, depuis la conquête normande, un nombre considérable de mots au français et au latin ; le vocabulaire d'une langue peut ainsi se renouveler presque totalement par des emprunts à des langues étrangères. Certains linguistes parlent alors de langues mixtes. L'expression est impropre. Car elle éveille l'idée qu'une pareille langue résulterait du mélange de deux langues placées dans des conditions *égales* et qu'on ne pourrait pas dire si une langue est la continuation d'une langue A ou d'une langue B, si le persan par exemple est la continuation de l'ancien perse ou de l'arabe. Dans les cas qui ont pu être observés d'une façon précise, on ne constate rien de pareil, et ce n'est pas un hasard historique ; cela tient aux conditions mêmes d'existence du langage. Chaque langue constitue un système, et les sujets bilingues qui ont le choix entre deux langues ne mêlent pas ces deux langues : quel que soit le nombre de mots arabes qu'il introduit, le Persan sait s'il adopte le système persan ou le système arabe, et il n'y a jamais de doute sur le système adopté. Les Anglais, qui ont admis des mots français innombrables, n'ont pas mélangé pour cela le système français au système saxon. Il y a donc, même dans les cas où les emprunts sont le plus nombreux, deux situations distinctes : celle de la langue que le sujet veut parler et dont il emploie ou cherche à employer le système linguistique, et celle à laquelle il emprunte d'autres éléments, dont le nombre peut d'ailleurs être aussi grand qu'on le voudra. Il y a, d'une part, un *fonds indigène*, et, de l'autre, des *emprunts*.

Une langue est définie par trois choses : un système phonétique, un système morphologique et un vocabulaire, c'est-à-dire

par une manière de prononcer, par une grammaire et par certaines manières de désigner les notions. Chacune de ces trois choses comporte un nombre pratiquement illimité de combinaisons possibles, et par suite on n'a aucune chance de trouver deux langues qui, distinctes par ailleurs, aient ou exactement le même système phonétique ou exactement le même système morphologique.

La prononciation et la grammaire forment des systèmes fermés ; toutes les parties de chacun de ces systèmes sont liées les unes aux autres. Le système phonétique et le système morphologique se prêtent donc peu à recevoir « des emprunts ». En fait il est rare qu'on emprunte à une autre langue soit un phonème (un son du langage), soit une forme grammaticale ; quand pareil fait se produit, il ne modifie pas l'ensemble de chacun des systèmes et demeure un accident. Au contraire, les mots ne constituent pas un système ; tout au plus forment-ils de petits groupes ; on peut soit changer le nom d'un objet, soit introduire un nom nouveau sans que cela retentisse sur l'ensemble du vocabulaire ; chaque mot existe pour ainsi dire isolément. Aussi peut-on emprunter à des langues étrangères autant de mots que l'on veut ; il suffit qu'une langue étrangère ait un prestige, qu'elle soit par exemple celle des maîtres comme l'a été l'anglo-normand en Angleterre à partir du XI^e^ siècle, pour que les emprunts se multiplient. Tous les mots ne s'empruntent pas avec une égale facilité : les mots usuels sont plus malaisément remplacés par des mots étrangers que les mots d'emploi rare ; on emprunte volontiers les noms d'objets nouveaux ou d'objets à la mode. La grammaire rend parfois les emprunts malaisés : une langue qui, comme le français, a des substantifs sans flexion, mais une conjugaison compliquée, emprunte volontiers des substantifs, mais relativement peu de verbes. En somme, le vocabulaire est le domaine de l' « emprunt ». Il n'y a pas de mot d'une langue dont on puisse dire *a priori* qu'il n'est pas emprunté à une langue étrangère. Il n'y a pas de langue ayant une histoire connue dont le vocabulaire ne renferme des emprunts. C'est donc avant tout par la persistance de la prononciation et de la gram-

maire que se traduit linguistiquement la volonté continue de parler une certaine langue qui définit la « parenté de langues ».

Ces systèmes se transforment souvent d'une manière complète. Notamment quand une langue est adoptée par une population qui perd son idiome indigène, le changement peut se précipiter ; la prononciation peut alors changer d'un coup ; la grammaire peut être simplifiée à l'extrême, comme il est arrivé dans les parlers créoles. Ici encore, certains linguistes seront tentés de parler de langues mixtes ; mais le *matériel* de la langue appartient à un idiome défini ; le créole de la Réunion ou de la Martinique est du français imparfait, mais c'est du français ; car c'est à l'imitation seule du français de leurs maîtres que les nègres l'ont constitué. La plus grande partie de la conjugaison a été sacrifiée ; mais ce qui en subsiste, l'infinitif, est français, et l'on n'y trouve pas le moindre élément africain. Il y a eu perte brusque d'une très grande partie d'un système grammatical au moment où une population de langue très différente et placée dans une situation sociale inférieure a appris une langue nouvelle. Les changements portent sur l'ensemble de chaque système, et l'on peut relier le système ancien au nouveau par un ensemble de formules de transformation. D'ailleurs, au début du moins, il subsiste toujours dans le système nouveau une portion notable du système ancien, et le peu que le créole a de grammaire est de la grammaire française. Tout ce qu'exige la définition, c'est qu'entre deux moments du développement de la langue immédiatement voisins l'un de l'autre, il y ait continuité dans l'emploi du système grammatical.

On ne peut pas dire que tout emprunt phonétique ou grammatical soit impossible. Dans certaines conditions favorables où des populations se mélangent d'une manière intime et où des sujets parlants, perdant le sentiment net d'appartenir à une certaine nation, ne s'attachent pas à parler leur langue avec pureté, il peut se produire des emprunts de ce genre. Ainsi, sur le domaine gallo-roman, à l'époque des invasions franques et de la domination mérovingienne, la langue de civilisation qu'était le latin s'est maintenue et a fini par s'imposer au peuple conqué-

rant, de langue germanique. Toutefois durant longtemps les conquérants ont gardé leur langue qui avait le prestige d'être la langue des maîtres. Le gallo-roman a donc emprunté beaucoup de mots à la langue des Francs. Mais cette langue comprenait des phonèmes dont le gallo-roman n'avait pas l'équivalent ; le latin vulgaire n'avait pas d'*h*, alors que *h* était fréquent en germanique, et le latin vulgaire avait un *v* labio-dental (le *v* français), alors que les Francs employaient la bilabiale *w* (*w* anglais). Les Francs qui parlaient latin ont gardé *h* et *w* dans les mots germaniques qu'ils mêlaient à leur latin, et les populations gallo-romanes ont tenu à prononcer correctement ces mêmes mots germaniques ; *hapia* a donc gardé son *h* dans la forme empruntée, et le mot est représenté en français par *hache*, avec une *h* « aspirée » ; *werra* a gardé son *w* qui a fini par se prononcer *gu*, et ce mot est représenté en français par *guerre* ; deux phonèmes nouveaux se sont introduits ainsi en gallo-roman. Ils ont même été adaptés à des mots latins ; *altu* « haut » du latin vulgaire a reçu *h* sous l'influence du synonyme germanique *hauh* ; *vespa* « guêpe » a reçu *w* sous l'influence du synonyme germanique *wefsa* ; et ainsi de nombreux autres cas : ceci montre que, pendant un temps, il a été élégant de prononcer ces mots latins avec un « accent » germanique. Des mots latins désignant des choses militaires se sont ainsi conservés avec une prononciation germanisée : *vagina* « fourreau » est devenu *wagina*, qui est représenté en français par *gaîne*. Ces deux additions au système phonétique latin sont quelque chose d'exceptionnel et dont les langues romanes autres que le français n'offrent pas l'équivalent ; chose curieuse, elles n'ont guère été durables : le *w* a rapidement passé à *gu*, puis à *g* tel qu'on l'avait dans des mots indigènes comme *goût* ; quant à l'*h*, elle s'est maintenue davantage ; mais elle a fini par disparaître aussi, et l'on sait que le français actuel n'a pas d'*h* ; la présence ancienne de *h* se traduit seulement par des hiatus : *la haute montagne*. Il y a aussi des emprunts grammaticaux ; mais, comme les emprunts de phonèmes qu'on vient de voir, ils sont liés à des emprunts de mots, et ils concernent ce qu'il y a pour ainsi dire de moins grammatical dans la grammaire.

Il n'y a pas d'exemple qu'une flexion comme celle de *j'aimais, nous aimions* ait passé d'une langue à une autre ; on n'emprunte une chose de ce genre que si l'on emprunte tout le système d'un coup, c'est-à-dire si l'on change de langue. Mais le turc a pu emprunter au persan le relatif *ki*. Prenant au latin des noms en *-arius* indiquant un métier, le germanique a pu développer un suffixe *-arja-*, et, de même que le latin avait *librarius* en face de *liber*, on a fait en gotique *bokareis* « scribe » en face de *boka* « lettre », *bokos* « livres » ; et ainsi le suffixe des noms d'agents en germanique (celui de l'allemand *schreiber*) se trouve être d'origine latine ; mais ce n'est encore que de l'emprunt de vocabulaire. On a même signalé, dans un endroit où le portugais et l'anglais étaient en contact, l'emploi de tours tels que *governador's casa* d'après l'anglais *governor's house* ; il y a eu ici emprunt de l'ensemble d'une expression ; d'ailleurs l'élément *'s* est si autonome dans ce tour anglais qu'il peut presque passer pour un mot, et le cas est assez comparable à celui de l'emprunt du relatif *ki* par le turc. Il arrive — en une mesure du reste assez faible et dans des situations très particulières — qu'on emprunte à une langue étrangère des petits mots à valeur grammaticale ; on n'emprunte guère de vraies formes grammaticales. Ainsi l'on est toujours ramené à la même conclusion : ce qui s'emprunte, ce sont essentiellement des éléments de vocabulaire.

L'exposé précédent repose sur l'idée que le sujet parlant a toujours l'intention de parler une langue définie ; et le principe est valable tant qu'il s'agit de langues nettement distinctes, qui excluent une communication entre sujets de langues différentes, comme le français, le provençal, l'italien, l'allemand, le hongrois, etc. ; peu importe que ces langues soient ou non de même famille. Les choses se présentent autrement si l'on envisage des parlers divers appartenant à un même groupe dialectal ou des patois employés par des gens qui ont par ailleurs une grande langue commune de civilisation de type voisin de leur parler local. Alors les sujets parlants peuvent imiter un parler qui passe pour plus élégant ; le français d'un paysan du centre de la France est du patois plus ou moins francisé, et toujours incom-

plètement; le patois de ce même paysan est souvent en grande partie du français patoisé. Il peut en pareil cas n'y avoir aucun sentiment net qu'on emploie tel ou tel parler ; alors il y a vraiment mélange, non de langues, mais de parlers divers d'une même langue, et l'on ne saurait toujours dire, au terme d'un développement linguistique de ce genre, quel parler a triomphé. Mais ceci n'empêche pas de savoir quelle langue ont parlée les sujets considérés; le français commun d'aujourd'hui repose avant tout sur le parisien; et il s'y mêle de plus en plus tant d'éléments provinciaux qu'il n'est pas licite de dire que ce soit du parisien ; en revanche, c'est du français, parce que les sujets n'ont jamais eu l'intention de parler autre chose que français ; on ne saurait dire au juste si le grec moderne est de l'attique ou de l'ionien ; car les sujets n'ont pas toujours voulu parler attique ou ionien ; mais il est évident que c'est du grec ionien-attique : la volonté de parler ce grec a été constante chez les sujets qui ont transmis la langue, et ceci suffit.

La définition de la parenté de langues ne s'applique donc qu'à de grands groupes, nettement distincts les uns des autres, non à des parlers distingués par de simples nuances. Et ceci encore résulte du caractère social de la définition.

II

Ces principes une fois posés, on voit comment peut se prouver une parenté de langues.

Partout où le système phonétique et le système grammatical présentent des concordances précises, où des correspondances régulières permettent de reconnaître l'unité d'origine des mots et du système phonétique et où le système des formes grammaticales s'explique en partant d'un original commun, la parenté est évidente. Les langues romanes sont visiblement parentes, parce que le nombre des concordances de détail qu'elles offrent est grand et que leur grammaire et leur système phonétique s'expliquent dans l'ensemble par une origine commune. L'identité des grammaires est évidente. Ainsi l'on a à l'infinitif : ita-

lien *cantare*, espagnol *cantar*, français *chanter* ; à l'indicatif présent, 3e personne, au singulier it. et esp. *canta*, fr. *chante*, au pluriel, it. *cantano*, esp. *cantan*, fr. *chantent* ; au prétérit simple, 2e personne, singulier it. *cantasti*, esp. *cantaste*, fr. *chantas*, pluriel it. *cantaste*, esp. *cantasteis*, fr. *chantates* etc. Ainsi que l'a bien montré, ici-même [1], M. Grammont, la constatation de la parenté facilite beaucoup l'apprentissage des langues en pareil cas : la seule correspondance, espagnol *h* = italien et français *f*, fait prévoir la forme de beaucoup de mots, comme esp. *haba* = ital. *fava*, fr. *fève* ; esp. *hacer* = ital. *facere*, fr. *faire* ; etc. Du même coup l'on voit que l'*h* espagnole, qui apparaît comme une divergence dans le système phonétique des langues romanes, est la transformation d'une *f*, et comment, par suite, l'espagnol se ramène historiquement au type général de ces langues. Il se trouve de plus que les vocabulaires concordent dans une large mesure ; le fait que l'espagnol a un grand nombre de mots arabes qui ne se retrouvent ni en italien ni en français ou que le français a des mots gaulois et germaniques qui ne se retrouvent ni en italien ni en espagnol ne change rien à la parenté ; car il s'agit d'emprunts. La preuve d'une parenté de langues est d'autant plus solide qu'elle se rapproche plus du cas qui vient d'être sommairement décrit.

Il va de soi que, pour établir une parenté de langues, il faut faire abstraction de tout ce qui s'explique par des conditions générales, communes à l'ensemble des langues. Ainsi les pronoms doivent être des mots courts, nettement constitués avec des éléments phonétiques aisés à prononcer, et en général sans groupes de consonnes. Il en résulte que les pronoms se ressemblent plus ou moins partout, sans que ceci implique une communauté d'origine. Et, d'autre part, les pronoms se ressemblent souvent assez peu dans des langues d'ailleurs très semblables ; qu'on compare par exemple les pronoms de l'arménien avec ceux du gotique ou de l'irlandais. Même des formes qui, en fait, se ramènent à un même prototype, comme *nous* du français et *us* de l'anglais,

1. *Scientia*, vol. XII, nº XXIV-4 (1912), p. 72 et suiv.

peuvent ne plus offrir un seul élément commun (l'*s* du français *nous* n'étant que graphique). On ne peut donc tirer parti des pronoms dans la détermination des parentés de langues qu'avec précaution.

Dans cet exposé on n'a tenu aucun compte des ressemblances générales de structure que peuvent présenter les langues considérées : le français et l'anglais s'accordent à caractériser plusieurs catégories grammaticales par un ordre de mots défini, à employer un article devant les substantifs, à se servir de petits mots accessoires pour indiquer la possession, l'attribution ; tout cela n'implique pas une origine commune, parce que, d'une part, les ordres de mots en question offrent trop peu de variations possibles pour que les concordances soient significatives, et que, d'autre part, les petits mots accessoires de même valeur grammaticale qui sont employés dans ces deux langues ne sont pas réductibles à des origines communes : ainsi angl. *the* et fr. *le*, angl. *of* et fr. *de*, angl. *to* et fr. *à*, angl. *we* et fr. *nous*, angl. *you* et fr. *vous*, angl. *he* et fr. *il* n'ont rien de commun. Si l'on n'avait pas les anciens dialectes germaniques d'un côté, et le latin de l'autre, la parenté du français et de l'anglais ne serait pas démontrable. Le chinois et telle langue du Soudan, celle du Dahomey ou ewe, par exemple, peuvent se servir également de mots courts, en général monosyllabiques, faire varier la signification des mots en changeant l'intonation, fonder leur grammaire sur l'ordre des mots et sur l'emploi de mots accessoires ; il n'en résulte pas que le chinois et l'ewe soient des langues parentes ; car le détail concret des formes ne concorde pas ; or, seule la concordance des procédés matériels d'expression est probante. C'est pour cette raison que les linguistes qui exigent des preuves rigoureuses ne considèrent pas comme établie l'existence de la famille ouralo-altaïque à laquelle on a autrefois attribué tant d'importance ; entre le turc d'une part, le finnois et le hongrois de l'autre, il y a des concordances générales de structure grammaticale ; mais ce qui prouve une parenté, c'est la concordance dans le détail matériel des moyens d'expression ; or, si l'on trouve des formations définies communes au finnois et au hongrois et si le finnois et le

hongrois se laissent expliquer par une même langue commune (qu'on voie par exemple le petit livre de M. Szinnyei, *Finnisch-ugrische Sprachwissenschaft*), on n'a constaté rien de pareil entre le finno-ougrien et le turc, et dès lors aucune parenté n'est démontrée. Le fait de procéder uniquement par suffixation, l'emploi de l'harmonie vocalique, etc. ne constituent pas des preuves de parenté.

Quand, dans son article de « Anthropos », VIII (1913), p. 389 et suiv., intitulé *The Determination of Linguistic Relationship*, un américaniste éminent, M. Krœber, a protesté contre l'emploi des concordances générales de structure morphologique pour établir des parentés de langues, il a eu entièrement raison. Seulement il n'est pas licite de conclure de là que les parentés doivent s'établir par la considération du vocabulaire, non par celle de la morphologie ; si juste qu'elle soit, la critique de M. Krœber ne justifie pas le procédé de certains américanistes qui fondent sur de pures concordances de vocabulaire leurs affirmations relatives à la parenté de telles ou telles langues entre elles. Les concordances grammaticales prouvent, et elles seules prouvent rigoureusement, mais à condition qu'on se serve du détail matériel des formes et qu'on établisse que certaines formes grammaticales particulières employées dans les langues considérées remontent à une origine commune. Les concordances de vocabulaire ne prouvent jamais d'une manière absolue, parce qu'on ne peut jamais affirmer qu'elles ne s'expliquent pas par des emprunts. On sait maintenant que les nombreux mots turcs présentés par le hongrois n'apportent à l'hypothèse d'une parenté du turc avec le hongrois aucun commencement de preuve. A en juger par le vocabulaire, l'anglais serait un mélange de germanique et de roman ; tout au plus pourrait-on constater que le fonds principal du vocabulaire courant, pronoms personnels, noms de nombre, noms de parenté, noms des parties du corps, noms des animaux les plus connus, verbes usuels tels que « manger » et « dormir », est germanique, et non roman. Mais il suffit de considérer la grammaire pour lever tout doute : rien dans le détail matériel de la grammaire anglaise ne

s'explique par le latin, tout s'y explique par la grammaire ancienne du germanique.

Il est vrai que la structure générale de la morphologie anglaise moderne diffère du tout au tout de celle de l'ancienne morphologie germanique ; mais ceci est sans importance : *la parenté n'implique aucune ressemblance actuelle des langues considérées*, ni surtout du système général des langues considérées ; et inversement il y a beaucoup de ressemblances, soit de structure générale, soit de vocabulaire, qui n'impliquent pas parenté.

Dans l'article cité ci-dessus, M. Krœber insiste sur l'importance qu'il y a à tenir compte du voisinage géographique des langues. Sans doute il arrive le plus souvent que les langues parentes occupent des aires contiguës ou du moins voisines. Mais, une fois mis à part ce fait grossier, il faut reconnaître que la contiguïté apporte à la démonstration linguistique de la parenté une gêne plutôt qu'un secours : les langues voisines sont celles qui ont subi les mêmes influences, qui ont emprunté les unes aux autres ou fait les mêmes emprunts à d'autres langues. La contiguïté des langues oblige donc à faire un départ souvent très délicat entre les emprunts et le vieux fonds de la langue, qui seul prouve en matière de parenté. En revanche, le grand éloignement géographique n'a pas empêché les linguistes de montrer que la langue de Madagascar représente la même langue ancienne que celles de Bornéo, de Java et des Philippines.

S'il s'agit de langues parentes qui ont cessé d'être très semblables entre elles, une forte ressemblance extérieure entre deux mots est, pour le linguiste, une raison de douter qu'ils représentent un même terme de la langue originelle. Le français *feu* n'a rien de commun avec l'allemand *feuer* ; en revanche l'allemand *feuer* remonte à un original que représentent par ailleurs des mots assez différents : grec *púr*, arménien *hur*. A qui sait que le latin, le germanique, le slave et l'arménien représentent une même langue commune, relativement ancienne, il est facile de montrer quels rapports il y a entre le français *cinq* (représentant le latin *quinque*) et les équivalents *five* de l'anglais, *piat'* du russe, *hing* de l'arménien ; mais, au premier abord, ces mots ne se

ressemblent guère. Ce n'est pas sur la ressemblance extérieure des mots que se fondent les linguistes pour les rapprocher et en faire l'étymologie, mais sur des formules de correspondances régulières. Ainsi à un *p* du grec ou du slave à l'initiale des mots, répondent *h* en arménien, *f* en germanique, l'absence de toute consonne en celtique. Ces formules une fois établies permettent de reconnaître dans ses grandes lignes l'histoire de la prononciation et d'établir comment les systèmes phonétiques se sont succédé les uns aux autres dans des langues apparentées.

Comme on l'a noté déjà, il n'y a pas lieu pour faire la preuve, d'exiger que toutes les formes grammaticales s'expliquent ; il suffit d'établir que des portions notables de la morphologie ancienne subsistent dans la langue considérée. Nulle part il n'est aussi aisé d'établir une parenté de langues qu'entre les langues indo-européennes, parce que la langue commune sur laquelle reposent les idiomes de cette famille comportait une morphologie très compliquée, d'une complication qui passe beaucoup la normale, et que de nombreux restes de ses formes ont subsisté dans chaque langue ; ce sont par exemple les verbes irréguliers du grec, les verbes forts du germanique, etc. Partout où l'on rencontre des débris importants du système verbal indo-européen, système que l'on connaît par les plus anciens textes, ceux de l'indo-iranien ancien et du grec ancien, on est sûr d'être en présence d'une langue indo-européenne. La démonstration de parenté est parfaite si l'on peut expliquer par la transformation des mêmes éléments anciens l'ensemble du système grammatical de deux langues distinctes.

A la longue cependant des langues parentes finissent par différer tant que leur communauté d'origine devient impossible à reconnaître. Si par exemple on n'avait que le français, le bulgare et l'arménien modernes pour représenter le groupe indo-européen, il ne serait pas aisé d'établir la parenté de ces trois langues, et l'on ne pourrait songer à en poser la grammaire comparée. Il suffit d'opérer avec ces mêmes langues, mais considérées sous des formes de quelques centaines d'années plus anciennes, à savoir le latin, le vieux slave des premières traductions et

l'arménien classique, pour que la parenté devienne évidente et pour qu'on puisse poser les principes essentiels d'une grammaire comparée de ces trois langues. La parenté de deux langues peut donc être, et est souvent, indémontrable, même alors qu'elle est réelle. On n'est jamais en droit d'affirmer que deux langues ne sont pas parentes au moins de loin : une parenté se découvrirait peut-être si l'on avait des formes plus anciennes de ces mêmes langues.

Bien qu'une démonstration complète de parenté de langues soit difficile à fournir, on a établi plusieurs grandes familles linguistiques : la famille indo-européenne qui occupe une aire immense et qu'on suit historiquement depuis près de trois mille ans ; la famille sémitique, moins vaste, mais dont les premiers textes écrits sont plus anciens ; la famille finno-ougrienne ; la famille bantoue ; la famille indonésienne ; la famille polynésienne (extraordinairement une) ; la famille dravidienne ; la famille caucasique du Sud ; d'autres encore. On voit que le principe ne s'applique pas seulement à l'Europe, ou aux langues indo-européennes, ou aux grandes langues de civilisation : il est universellement valable. Il ne manque pas de cas où le travail n'est pas encore fait, mais déjà l'on entrevoit clairement la possibilité de mettre en évidence certaines autres familles linguistiques définies : il doit y avoir par exemple une famille hamitique dont la grammaire comparée n'est pas encore faite. Il suffit d'un travail méthodique pour aboutir à coup sûr à des résultats certains sur plusieurs domaines.

Mais il y a de graves difficultés.

Tout d'abord une difficulté de fait. Pour beaucoup de langues de peuples non civilisés, on n'a que des vocabulaires, et la grammaire est ou inconnue, ou connue d'une manière toute partielle. Si, en pareil cas, on observe un très grand nombre de communautés de vocabulaire entre certaines langues, et si ces communautés concernent les mots les moins sujets à emprunt, notamment les verbes qui indiquent les actions usuelles comme *aller* et *venir*, *boire* et *manger*, *vivre* et *mourir*, *entendre* et *voir*, *dire* et *se taire*, etc., ou des adjectifs comme *vieux* et *neuf*, *grand*

et *petit*, *long* et *large*, etc., ce serait pur pédantisme que de se refuser à en faire usage. Seulement il ne faut pas se faire illusion sur la rigueur de la preuve ainsi faite, bien que la possession en commun d'un certain fonds de vocabulaire indique le plus souvent une parenté. Là où l'on n'a pas d'autres données, on peut provisoirement, et en faisant les réserves nécessaires, se servir des indications ainsi obtenues. L'observation attentive du vocabulaire conduit du reste presque toujours en pareil cas à relever quelques coïncidences grammaticales qui achèvent la démonstration.

En second lieu, les langues spéciales échappent en partie à la définition qui résulte des considérations présentées ci-dessus. Elles n'y échappent pas à la lettre ; elles la confirment même en un sens : les langues spéciales ne comportent en général ni une prononciation, ni une grammaire propres; l'argot français par exemple est caractérisé uniquement par des mots particuliers. Et ceci montre comment le vocabulaire est indépendant des systèmes phonétique et morphologique qui définissent une langue : l'argot est une langue spéciale à l'intérieur du français. La chose va parfois très loin ; ainsi le tsigane arménien est purement de l'arménien pour la prononciation et la grammaire ; mais le vocabulaire n'a rien d'arménien ; ceci s'explique par le fait que les Tsiganes d'Arménie, sachant l'arménien, ont employé uniquement le système arménien, mais que, désirant parler une langue spéciale, inintelligible au reste de la population, ils ont gardé leur vocabulaire traditionnel. Si l'on applique à la lettre la définition, le tsigane arménien est purement et simplement de l'arménien ; mais il faut convenir qu'on est ici en présence d'un cas tout particulier. Ceci revient à dire que la définition générale des parentés de langues, faite pour des populations de type normal, s'applique mal aux langues spéciales de populations parasites. Toutefois, comme, en ce cas, il n'y a mélange ni au point de vue phonétique ni au point de vue grammatical, le principe est rigoureusement confirmé.

En troisième lieu, il y a des langues qui, étant entourées de langues d'une autre famille et paraissant être des débris isolés

d'une famille disparue, ne se laissent pas grouper avec les langues voisines et ne rappellent même aucune langue connue. Leur système grammatical n'offre de concordances précises avec aucun autre idiome, ou du moins n'en signale-t-on pas. C'est le cas du basque par exemple. Un linguiste illustre, un de ces hommes toujours rares dont les connaissances sont larges et qui se soucient plus de ce qui reste à trouver que de ce qui est déjà découvert, M. H. Schuchardt, a été conduit à rechercher si le basque n'offrirait pas avec des langues hamitiques quelques concordances; il y a relevé des mots qui rappellent des mots nubiens, et, d'une manière générale, des mots hamitiques, et a signalé ces concordances dans deux articles récents. Mais ces concordances sont vagues, peu nombreuses. Il n'est pas évident qu'on ne trouverait pas entre le basque et le vocabulaire d'un groupe quelconque de langues des concordances analogues. Dans son *Baskisch und Hamitisch*, pag. 6 (extrait de la « Revue internationale des études basques », VII (1913), M. Schuchardt dit que les jeunes gens dont le coup d'œil n'est pas assez formé sont trop portés à tenir les comparaisons de mots pour fortuites : on en juge tout autrement, dit-il, quand on s'est beaucoup exercé aux comparaisons de mots et qu'on a acquis en cette matière un certain sentiment statistique. Mais on ne peut pas fonder une doctrine sur un sentiment individuel. Et d'ailleurs, même si les ressemblances de vocabulaire signalées ne sont pas fortuites, on n'a aucun moyen de montrer qu'elles ne proviennent pas d'emprunts. Peut-être dira-t-on que, si l'on n'avait de langues indo-européennes que le français actuel et l'arménien moderne, les seules traces de la parenté de ces deux langues qu'on pourrait découvrir seraient des mots, comme le nom de *mère*, en arménien actuel *mer* ou *mar*, suivant les dialectes ; ou le nom de nombre *dix*, en arménien *tas* ou *das* suivant les dialectes ; et ce sont en effet des *traces* valables pour qui connaît la parenté des deux langues, mais ce ne seraient pas des *preuves* pour qui ne la connaîtrait pas, et en fait la ressemblance frappante de ces mots, effectivement parents, provient d'une série d'accidents fortuits. Au lieu de converger, ces mots auraient pu tout aussi bien diverger comme

fr. *père* et arm. *her* ou *har*, fr. *vingt* et arm. *khsan*, qui sont également parents ; et la communauté d'origine serait impossible à reconnaître directement. Si donc on peut d'abord constater des ressemblances de vocabulaire entre deux ou plusieurs langues pour indiquer de quel côté il faut chercher, ce n'est pas de là que peut venir une démonstration définitive ; le vocabulaire ne peut servir qu'à orienter la recherche; la preuve se trouve ailleurs.

Enfin, le procédé de démonstration des parentés de langues indiqué ici s'applique bien à des langues dont le type originel a comporté une grammaire compliquée. S'il s'agit de la famille indo-européenne ou de la famille bantoue, il n'y a pas de difficulté; tant qu'on n'a pas affaire à des langues séparées de la période commune par un trop long intervalle de temps ou par des altérations trop profondes, les choses sont même le plus souvent évidentes du premier coup. D'anciennes langues indo-européennes comme le gotique ou le slave ont une grammaire très différente de l'ancienne grammaire indo-européenne ; mais il subsiste assez de restes de cette grammaire, soit parmi les formes régulières, soit surtout parmi les formes anomales pour que le caractère indo-européen de la langue se voie sans exiger une démonstration. Si transformé, si éloigné de l'ancien type indo-européen que soit le « tokharien » récemment découvert en Asie centrale, on l'a du premier coup reconnu pour indo-européen ; le nombre des particularités grammaticales indo-européennes conservées y est encore grand. Mais, si l'on est en présence de langues qui n'ont presque pas de grammaire, si presque toute la grammaire proprement dite tient en quelques règles de position relative des mots, comme dans certaines langues d'Extrême-Orient ou du Soudan, le procédé ne s'applique pas. Et alors la question des parentés de langues est pratiquement insoluble, aussi longtemps qu'on n'aura pas trouvé de criteres qui permettent d'affirmer que des langues de ce type sont issues les unes des autres et que les ressemblances de vocabulaire qu'elles offrent ne sont pas dues à des emprunts. Il ne résulte pas de là que le principe, applicable ailleurs, devienne mauvais ici, mais seulement que certaines

langues ne comportent pas, en l'état présent des connaissances, une classification généalogique. Il appartient aux linguistes qui s'occupent de ces langues d'aviser à tourner la difficulté. La chose n'est pas impossible, pour peu qu'on sache rester fidèle à l'esprit du principe plutôt qu'à sa lettre.

Si même une solution ne se laissait pas trouver, il n'y aurait pas lieu de critiquer les linguistes à cause de cette impuissance : il est remarquable qu'on puisse en certains cas faire la preuve de la parenté de langues par des faits d'ordre purement linguistique ; mais il n'est pas surprenant que cette preuve ne soit pas toujours possible ; il est naturel qu'elle s'applique seulement aux cas où les faits envisagés ne sont pas séparés du début de l'époque proprement historique par un trop long espace de temps et où les transformations n'ont pas été trop profondes. De par sa nature, la classification généalogique des langues admet d'être incomplète. Elle se complétera au fur et à mesure qu'on acquerra des données sur l'histoire des langues.

C'est dire que le problème de l'unité d'origine des langues qui préoccupe certains linguistes ne saurait se discuter actuellement. Les problèmes qui peuvent être utilement posés à cet égard sont ceux qui sont relatifs à la constitution exacte de familles qu'on entrevoit, mais qu'on n'a pas étudiées avec précision. Les familles qu'on a constituées jusqu'ici sont celles qui apparaissent aisément ou celles qui renferment des langues de civilisation importantes. Le travail qui reste à faire est plus décourageant, car il s'agit en partie de langues où les relations de parenté sont malaisées à établir, en partie de langues sauvages peu connues et que presque personne n'a d'intérêt pratique à étudier. Il a été aisé d'établir la parenté de langues dont on avait de bonnes grammaires et dont la philologie était faite ; on est maintenant devant des langues dont il faut faire la grammaire et dont la plupart, faute de textes anciens, ne comportent aucune philologie. Il y a donc un grand travail d'observation à faire. Si des résultats plus nombreux ne sont pas acquis, cela tient avant tout au petit nombre ou à l'absence de savants qualifiés capables de faire ce travail. Pour compléter rapidement la classification généalogique des langues, il

suffirait de former quelques équipes de linguistes. Un seul bon travailleur peut suffire à poser une grammaire comparée là où l'on a de bonnes descriptions; c'est ainsi que M. Brandstetter constitue actuellement la grammaire comparée des langues indonésiennes.

Il n'y a pas de raison de croire que certaines langues communes d'où sortent les grandes familles établies ne sont pas, à leur tour, des transformations d'une même langue plus ancienne. On peut imaginer, par exemple, que l'indo-européen, le sémitique (qui est apparenté au hamitique), le caucasique du Sud et le finno-ougrien sortiraient d'une même langue plus ancienne; il a déjà été signalé des concordances entre l'indo-européen et le sémitique, entre l'indo-européen et le finno-ougrien, entre le caucasique du Sud et le sémitique. Mais beaucoup de ces concordances ne concernent que le vocabulaire et ne sauraient passer pour probantes; celles qui concernent la grammaire sont peu nettes et peu nombreuses. La preuve n'est pas faite; elle ne pourrait du reste être acquise que le jour où la grammaire comparée du sémitique et du hamitique sera constituée, ainsi que celle du caucasique du Sud dont il ne faut peut-être pas séparer les langues caucasiques du Nord, lesquelles sont très variées; il faudrait te compte des langues d'Asie Mineure, du lycien qu'on ne comprend guère, du hittite qu'on n'a pas encore déchiffré [1], ou de langues plus lointaines comme l'élamite. Il se pose là de grands problèmes, dont la solution n'est pas désespérée. Mais on a souvent eu le tort, dans les derniers temps, de vouloir les résoudre avant d'avoir fait les travaux préparatoires nécessaires. Tant que la grammaire comparée du hamitique et celle du caucasique seront à peine esquissées, tant que, par suite, la position exacte du sémitique sera inconnue, il sera prématuré de vouloir relier le sémitique à l'indo-européen.

1. Le présent article a paru en 1914; les textes hittites sont maintenant lus et déchiffrés, surtout par M. Hrozný; mais l'interprétation est encore très hypothétique, et l'affirmation que le hittite serait indo-européen paraît bien aventurée; elle a été contestée par la plupart de ceux qui ont examiné les documents.

Toutefois les meilleurs linguistes feront œuvre vaine s'ils s'attaquent directement à des langues trop différentes ; on pourra peut-être rapprocher un jour les grammaires comparées de l'indo-européen, du sémitique, du hamitique, du caucasique, du finno-ougrien ; il y a là une hypothèse qu'on pourra essayer de vérifier quand les travaux préparatoires suffisants seront faits, et que, dès maintenant, plus d'un fait rend vraisemblable. Mais il serait vain de vouloir comparer aujourd'hui le latin, l'hébreu et le géorgien. Tout essai de ce genre est une faute évidente contre la méthode.

La classification généalogique des langues est en fait, on le voit, chose différente des classifications biologiques. Une classification telle que la classification linnéenne consiste à rapprocher des êtres qui ont des structures anatomiques analogues et où au moins le commencement du développement de chaque individu depuis la cellule initiale est comparable, donc des êtres qui, dans leur développement individuel, offrent *actuellement* des caractères communs, alors même que, à l'état adulte, leur aspect extérieur et les fonctions remplies par certains de leurs organes diffèrent le plus.

Il est possible, et l'on admet en général, que les êtres qui sont ainsi rangés dans une même classe sortent d'un même ancêtre ou d'ancêtres exactement semblables ; mais cette hypothèse, qui explique les ressemblances constatées, n'est pas le principe de la classification, bien que les biologistes tendent de plus en plus à tenir compte de la façon dont ont évolué les êtres qu'ils étudient et qu'ils classent, et aussi à ne pas négliger les rapports entre les diverses classes. Le jour où les biologistes arriveraient à suivre exactement l'évolution qui a abouti à différencier les espèces, leurs classifications deviendraient comparables aux classifications linguistiques ; entre les unes et les autres, la différence essentielle serait seulement celle du degré de rapidité de l'évolution, l'évolution linguistique étant beaucoup plus rapide, l'évolution biologique souvent si lente que les espèces semblent stables.

Pour le linguiste, seul le fait historique de la continuité entre une langue ancienne et des langues postérieures entre en considération. En l'état présent des choses, l'anglais et le russe sont

deux langues de type absolument distinct, et c'est à peine si l'on y peut discerner quelques éléments de vocabulaire semblables ; cela n'empêche pas que ce soient deux langues indo-européennes ; la preuve résulte de ce que l'anglais moderne continue le vieil anglais, et le russe moderne, le vieux russe ; or, le rapprochement du vieil anglais et du vieux russe est démontrable directement, et surtout il est facile de prouver que le groupe germanique dont l'anglais fait partie et le groupe slave dont le russe est l'un des représentants sont tous deux des formes prises par la langue indo-européenne commune. Tant que deux langues de même famille subsistent, leur appartenance à cette famille ne saurait subir un changement, quelles que soient les transformations subies, et même si les altérations intervenues ne laissent subsister aucune trace de la commune origine des deux idiomes.

La détermination de la famille à laquelle appartient une langue est une donnée indispensable à fixer pour faire l'histoire de cette langue. Mais, cette donnée une fois acquise, il reste à déterminer le détail des changements intervenus et à reconnaître les influences subies. Ces changements et ces influences sont souvent pour beaucoup plus que la langue initiale dans le résultat final. La notion de parenté de langues est chose précise ; mais, justement parce qu'elle est définie d'une manière précise, elle n'est que l'une des données avec lesquelles opère l'historien du langage.

Comme la parenté de langues est rigoureusement déterminable et que, au contraire, à part les emprunts de mots, les influences qui déterminent les changements linguistiques ne se laissent ni reconnaître d'une manière exacte, ni surtout établir d'une manière certaine, la parenté tient, dans les théories des linguistes, une place qui dépasse — et sans doute de beaucoup — son importance réelle. Mais, en la définissant bien et en n'oubliant pas que la tradition continue des parlers n'est ni le tout ni toujours le principal de ce qui conditionne le développement des langues, on peut fonder sur la théorie de la parenté la théorie historique des langues.

LES PARENTÉS DE LANGUES [1]

M. Schuchardt a précisé ses vues sur les parentés de langues dans deux mémoires, l'un daté de juillet 1914, dans *Nordisk Tidsskrift for Filologie*, IV, 6, p. 145-151, l'autre dans *Sitzungsberichte* de l'Académie de Berlin, XXXVII (1917), p. 518-529. Ces mémoires se réfèrent à celui que j'ai publié dans *Scientia*, XV (1914), p. 403 et suiv. (voir ci-dessus, p. 76 et suiv.), et où il y avait un renvoi aux articles publiés sur la question par M. Schuchardt, dans la *Revue basque*.

Il est inutile de le dire, les vues de M. Schuchardt sur la parenté des langues sont d'un vif intérêt. M. Schuchardt n'est pas seulement l'un des plus grands romanistes de ce temps, il est aussi l'un des défricheurs les plus hardis de domaines linguistiques nouveaux : caucasique, basque, hamitique. Parmi les linguistes d'aujourd'hui, il n'en est pas de plus personnel.

Je souscrirais d'ailleurs à beaucoup des affirmations de M. Schuchardt.

Comme toutes les expressions figurées employées en linguistique, l'expression *parenté* de langues est trompeuse : la parenté de langues est autre chose que ce que l'on appelle d'ordinaire parenté ; une langue « fille » est une transformation d'une langue « mère », et non un rejeton. L'expression est trop établie pour qu'on y renonce ; il suffit de la définir pour n'en être pas dupe.

Dans la grammaire comparée actuelle, les langues communes tiennent une grande place : à l'indo-européen commun, on superpose un italo-celtique, à l'italo-celtique, un italique et un celtique, au celtique et à l'italique, un gaélique et un brittonique d'une part, un latin et un osco-ombrien de l'autre, etc. Et, dès que deux faits paraissent concorder dans des langues ainsi groupées, on est disposé à en reporter l'origine à la période de communauté, pour peu qu'on n'ait pas la preuve d'une date plus

1. Extrait du *Bulletin de la Société de linguistique*, XXI (n° 66), p. 9 et suiv.

récente. M. Schuchardt, qui a été le premier à reconnaître la continuité des aires linguistiques et à écarter la notion vulgaire de dialectes ayant des limites arrêtées, critique cette grave faute de méthode. Une langue commune, une *Ursprache* comme on dit en allemand, ne se constitue que dans certaines conditions, et pour en prouver l'existence, il faut avoir des concordances d'un type tout particulier. Il ne faut surtout pas attribuer à une ancienne communauté ce qui provient de développements parallèles, mais indépendants.

Une entité comme le « latin vulgaire » est chose fallacieuse. Les langues romanes continuent le latin, et toutes présentent, par rapport au latin classique, des innovations communes. Il y a donc un « roman commun » dont la définition résulte de la comparaison des langues romanes et dont, du reste, beaucoup de traits se trouvent attestés par des témoignages remontant à l'époque de l'empire romain. Mais c'est chimère que de réaliser les innovations du « roman commun » dans un « latin vulgaire » qui aurait été parlé on ne sait quand, et on ne sait par qui. Les traits linguistiques par lesquels le « roman commun » se distingue du latin classique, résultent de faits complexes et dont l'analyse est malaisée.

Ceci dit, il faut examiner l'idée fondamentale de M. Schuchardt.

Voici le fait à interpréter : toute langue comprend, en proportions variées, des éléments qui proviennent de plusieurs langues différentes.

On admet d'ordinaire que ces éléments doivent être considérés de deux manières. Les uns proviendraient de la langue dont la langue considérée est la continuation, et les autres de langues étrangères : les premiers sont les éléments *indigènes* ; aux seconds, on donne le nom d'*emprunts*. Si donc on envisage l'histoire d'une langue entre deux dates données, les éléments *indigènes* sont ceux qui se sont transmis sans interruption entre les deux dates considérées ou qui ont été faits avec des éléments indigènes ; les *emprunts* sont les éléments, pris à des parlers quelconques, et qui ne reposent pas sur une tradition continue.

Cette doctrine suppose que les sujets qui ont transmis les éléments indigènes ont eu constamment, d'une manière plus ou moins nette, le sentiment et la volonté de parler leur langue traditionnelle. Quel que soit le nombre, quelle que soit l'importance des « emprunts », il y a des éléments *indigènes* là où les sujets ont toujours cru et voulu parler une langue définie. Par exemple, quelle que soit en anglais la part de l'élément français, les sujets anglais ont toujours eu le sentiment et la volonté de parler leur langue nationale, et non celle des barons franco-normands. Ce qui importe, ce n'est pas de déterminer la proportion de tel ou tel élément, mais de savoir quelle langue ont cru et voulu parler ceux qui ont fait la transmission continue entre les deux dates considérées. M. Schuchardt ne paraît pas attribuer une importance décisive à cette considération que j'ai introduite, et sans laquelle la doctrine classique me paraît, comme à lui, théoriquement insoutenable.

M. Schuchardt ne se place pas au point de vue des sujets parlants, mais au point de vue de la langue. Il constate qu'il y a mélange (Mischung) : « Le mélange pénètre tout le développement linguistique ; il intervient entre langues distinctes, entre parlers proches, entre langues parentes et entre langues non parentes. Qu'il s'agisse de mélange ou d'emprunt, d'imitation, d'influence etrangère, nous sommes toujours en présence de phénomènes essentiellement semblables. »

Ce qui, en effet, pour M. Schuchardt, est essentiel, ce n'est pas le sentiment et la volonté de continuer telle ou telle langue ; c'est purement le souci d'être compris de ceux à qui l'on parle.

Qu'arrive-t-il en fait dans les groupes de langues bien observées ?

*
* *

Dans le domaine occupé par les langues indo-européennes, la question de savoir si une langue est ou non indo-européenne ne se pose jamais : la réponse est toujours évidente [1]. La question

1. Si en certains cas on émet des doutes, ainsi pour le hittite (voir par

de savoir si une langue appartient à tel ou tel des grands groupes de l'indo-européen, l'indo-iranien, le slave, le germanique, le latin, etc., ne se pose pas davantage : la réponse n'est pas moins évidente. C'est que les membres de la nation qui parlait l'indo-européen ont fait prévaloir une langue bien définie sur un vaste domaine. C'est que les membres de chacune des nations qui se sont constituées par la suite ont de même fait prévaloir sur des domaines étendus les formes nouvelles, également bien définies, prises parmi eux par l'indo-européen. Sans doute on a signalé çà et là des populations mixtes qui sont dans un état linguistique trouble. Mais ce ne sont pas ces populations qui l'emportent. Dans tous les cas clairs qu'on connaît, un groupe de langues indo-européennes résulte de l'extension d'une langue ayant une force d'expansion, et qui, par suite, appartient à une population ayant un sentiment national et la conscience de son individualité.

Chacun des groupes indo-européens représente une déviation spécifique de l'indo-européen commun, parce qu'il provient d'une nation à part qui avait donné à l'indo-européen un aspect nouveau, distinct de tout autre.

Donc, dans le groupe linguistique de tous le mieux étudié, le groupe indo-européen, le départ entre ce qui est indigène et ce qui est emprunté se fait nettement. On peut, pour tel ou tel détail de l'arménien, par exemple, se demander s'il est indigène ou emprunté au parthe ; mais la masse indigène et la masse parthe de l'arménien s'opposent l'une à l'autre ; on ne saurait les mettre sur un même plan : il s'agit de deux ordres de faits différents ; et, même pour le vocabulaire, la décision n'est presque jamais incertaine : on peut, un instant, se demander si le nom arménien du « bras », *bazuk*, est indigène ou emprunté ; mais on se rend aisément compte, par le suffixe *-k-*, qu'il doit être emprunté ; du reste, il n'y avait pas de nom indo-européen du « bras ». Les cas douteux sont négligeables, et du reste peu

exemple C. D. Buck, *Classical Philology*, XV [1920], p. 84 et suiv., et 203 et suiv.), ce n'est que pour des langues mal connues, et dont les textes ne sont pas interprétés d'une manière sûre.

intéressants au fond. La morphologie, le traitement phonétique des éléments traditionnels sont indigènes ; ce sont seulement certaines catégories du vocabulaire — très étendues à la vérité —, certaines manières de s'exprimer qui sont empruntées au parthe.

Le groupe sémitique, le groupe finno-ougrien, le groupe bantou se prêteraient à la même observation.

Ce que présente l'histoire des langues dans les cas bien observés et bien décrits, ce sont des extensions de langues définies parlées par des nations ayant conscience d'elles-mêmes. En fait, la distinction d'un élément indigène et d'un élément emprunté y est nettement tranchée.

Les parlers qui sont des mélanges informes de deux langues différentes comme le slavo-italien et l'italo-slave qu'a décrits M. Schuchardt sont ceux de populations inférieures ; ils ne survivent généralement pas. Au cas où ils survivraient, il est permis de se demander si l'on en pourrait faire la théorie : les faits seraient beaucoup trop compliqués. On se trouverait sans doute devant des parentés indéterminables.

Ce n'est pas à dire que, si l'on envisage le résultat final d'un développement, la langue indigène y soit pour plus ou même pour autant que les « emprunts ». Mais, dès l'instant que l'on définit une famille de langues par un sentiment et une volonté continus de parler une même langue, — ce qui répond à la réalité observée dans tous les cas connus —, ceci n'a aucun inconvénient pour la définition.

Entre deux moments éloignés du développement d'une seule et même langue, le type linguistique peut changer du tout au tout. La structure profonde de la langue, ce que M. Schuchardt appelle *die innere Form*, est autre en français ou en anglais qu'elle n'était en indo-européen. Je l'ai si peu méconnu que je l'ai dit expressément, mais en indiquant que cela n'a aucune importance pour la définition qui est donnée des familles de langues, puisque la parenté de langues, expression d'un fait historique, n'implique aucune communauté *actuelle*, si petite soit-elle, entre les langues considérées.

Le changement de structure peut tenir à des influences étran-

gères. Une population qui apprend une langue nouvelle tend souvent à introduire dans cette langue des changements plus graves qu'une population qui continue son usage ancien. Si, par exemple, l'arménien et l'iranien occidental (notamment le persan) ont perdu toute distinction de genre grammatical, cela peut tenir à ce que l'arménien et l'iranien occidental ont été adoptés par des populations parlant autrefois des langues — telles que les langues caucasiques du Sud ou l'élamite — auxquelles le genre grammatical était inconnu. Les tendances au changement peuvent ainsi procéder d'influences étrangères. Mais ceci ne change rien au fait fondamental que les sujets parlants ont voulu parler et ont parlé telle ou telle langue.

Faire une classification généalogique des langues n'aboutit pas à traiter en quantité négligeable les éléments non indigènes et les influences étrangères, mais à traiter les éléments *indigènes* et les éléments *empruntés* comme résultant de procès historiques spécifiquement distincts.

La différence tient au sentiment des sujets parlants. Mais elle se traduit par un fait linguistique. En pratique, on n' « emprunte » ni une forme grammaticale, ni un phonème : l'anglais qui a tant pris au franco-normand, l'arménien qui a tant pris au parthe, ne doivent, l'un au franco-normand, l'autre au parthe, ni une forme grammaticale, ni un phonème. Hors le cas de bilinguisme, l' « emprunt » porte, à peu près exclusivement, sur la partie de la langue qui, à la différence de la morphologie et de la prononciation, ne constitue pas un système fermé. C'est une conséquence de la définition. Et, ici encore, on n'est pas en présence d'une théorie, mais de faits positifs.

M. Schuchardt dit, avec raison, que la distinction entre le vocabulaire et la morphologie n'est pas absolue. Le fait que le pronom singulier de 2e personne est en latin *tū* est un fait de vocabulaire ; en français actuel, *tu* n'est plus un mot autonome ; ce n'est que la caractéristique de la 2e personne du singulier des verbes. Du latin *tū*, qui était un mot autonome, au français *tu*, qui est un pur élément grammatical, il y a eu glissement, et l'on ne peut marquer le moment où *tu* a cessé d'être un mot pour

devenir une caractéristique grammaticale ; il y a eu transition insensible d'une valeur à l'autre. Tel élément qui est morphologique peut donc être issu d'un mot emprunté ; il ne résulte pas de là qu'une forme grammaticale proprement dite soit empruntée.

Quand on n'a, pour rapprocher deux langues, que quelques ressemblances de mots, il est impossible de dire si ces ressemblances sont fortuites comme celles entre fr. *feu* et all. *feuer*, fr. *faillir* et all. *fallen*, fr. *gros* et all. *gross*, etc., si elles résultent d'emprunts, ou enfin si elles sont dues à une parenté proprement dite, du type défini ci-dessus.

Plus nettement les éléments avec lesquels on opère sont de caractère grammatical, et plus ils sont propres à prouver une parenté de langues. Plus nettement ils sont de purs faits de vocabulaire, et moins ils sont aptes à établir, *dans une mesure quelconque*, une parenté, au sens précis attribué à ce mot par la définition adoptée ci-dessus.

Il reste d'ailleurs vrai que, dans la plupart des cas, les mots courants du vocabulaire demeurent en notable partie les vieux mots traditionnels. Dans une langue aussi profondément altérée à tous égards que l'est l'anglais, l'examen des pronoms *me*, *we*, *us*, *you*, des noms de nombre *one*, *two*, *three*, *ten*, des noms de parenté *father*, *mother*, *brother*, *sister*, *son*, *daughter*, des verbes *be* (*is*, *was*), *come*, *eat*, *love*, *bear*, etc., suffirait à dénoncer le caractère indo-européen de l'anglais.

*
* *

Toutes les langues n'ont pas, comme l'indo-européen, le sémitique, le finno-ougrien, le bantou, des systèmes morphologiques nets et dont les continuations diverses sont aisément reconnaissables grâce au maintien de certaines parties de ce système. Dans des langues où il n'existe pas de système morphologique comportant des formes pourvues de caractéristiques singulières, la preuve d'une parenté peut être très difficile à administrer.

D'autre part, s'il s'agit de langues qui ont beaucoup divergé, soit par suite du long temps depuis lequel elles se sont séparées, soit par suite de la rapidité avec laquelle elles se sont transformées, une parenté réelle peut être devenue indémontrable parce que tous les faits morphologiques communs sont effacés.

La difficulté qu'on éprouve à démontrer une parenté de langues en bien des cas ne suppose donc pas qu'il s'agisse de langues mixtes. On ne saurait fonder une théorie linguistique générale sur des langues où le premier travail de défrichement commence seulement à se faire. Avant de conclure que la théorie classique des familles de langues ne s'applique pas aux cas obscurs, il faudra établir des grammaires comparées qui ne sont pas faites jusqu'ici : pour les domaines basque, hamitique et caucasique auxquels s'intéresse M. Schuchardt, l'étude comparative est à peine amorcée. La grammaire comparée du sémitique est moins précisément faite que celle de l'indo-européen ; celle du berbère existe à peine ; la grammaire historique de l'égyptien commence à se constituer ; les rapports des langues caucasiques entre elles ne sont pas établis, sans parler de ce que l'on pourra tirer des découvertes sur les anciennes langues d'Asie Mineure ; la grammaire comparée des parlers basques est tout au plus esquissée. Le caractère plus ou moins aberrant des vocabulaires n'a rien de décisif. Tant qu'on n'aura pas tiré de la comparaison des parlers de chacun des groupes et de l'examen des anciens textes tout ce que l'on en peut obtenir, les essais de rapprochement de vocabulaire entre ces divers groupes de langues ont peu de chance d'aboutir à des conclusions solides. Si intéressants qu'ils soient, et propres à fournir une première orientation, ils ne sauraient servir à fonder aucune théorie.

DIFFÉRENCIATION ET UNIFICATION DANS LES LANGUES [1]

On enseigne que le français et le bulgare, par exemple, sont des langues indo-européennes, le babylonien et l'arabe du Maroc des langues sémitiques, le souahéli et le douala des langues bantoues, le lapon et le magyar des langues finno-ougriennes, et ainsi de suite. On entend par là que, abstraction faite des différences que ces langues présentent maintenant entre elles, le français et le bulgare continuent une seule et même langue, dite langue indo-européenne; et de même pour les autres cas. Cette formule a un sens précis si l'on considère les langues, abstraction faite des hommes qui les parlent et des lieux où on les parle. Entre le français moderne et l'indo-européen, il y a une tradition continue, et du moment le plus ancien au moment actuel, des groupes d'hommes se sont trouvés qui employaient, autant qu'il leur était possible, le même système phonique et le même système morphologique, la même prononciation et la même grammaire. Le point de vue auquel on se place pour faire une classification généalogique des langues est donc légitime, et l'on ne saurait contester la correction du procédé. Mais la transmission des systèmes considérés se fait de manières distinctes suivant les cas. Et, constamment, deux tendances antagonistes sont en jeu, l'une vers la différenciation, l'autre vers l'unification.

Par le fait qu'elle est employée, toute langue tend à se différencier de plusieurs manières.

Au moment où il apprend à parler, l'enfant ne reçoit pas le

1. *Scientia* (*Rivista di scienza*), vol. IX, V (1911), XVIII, 2.

langage tout fait, on le sait; il doit, à l'imitation des personnes qu'il a occasion d'observer, se créer un système d'articulations, un système grammatical et un vocabulaire parallèles à ceux de son entourage. Sauf le cas — du reste fréquent — des enfants anormaux, cette imitation réussit d'une façon sensiblement exacte; mais elle ne saurait être parfaite. Entre le langage que s'est fixé l'enfant, une fois achevée la période de l'apprentissage de la langue, et celui des personnes qu'il a imitées, il y a des différences appréciables. On constate en fait que, pour une part, ces différences sont les mêmes chez les enfants nés vers le même temps, en une même localité, dans un même milieu social. Il n'y a pas lieu de chercher ici en quelle mesure les innovations communes à toute une génération résultent de l'état de la langue au moment où elles ont lieu, et en quelle mesure elles peuvent être dues à des tendances héréditaires existant dans une population. Il suffit de noter que, par le fait même qu'elles sont communes, ces innovations subsistent, qu'elles sont reproduites par les générations ultérieures et peuvent éventuellement servir de points de départ à d'autres innovations. D'autre part, l'emploi qui est fait du langage amène les adultes à transformer leur usage : les mots souvent rapprochés tendent à se souder les uns aux autres, les formes souvent employées perdent leur valeur expressive et se prononcent d'une manière plus rapide et plus sommaire; le détail des faits de ce genre varie d'une localité à l'autre. A moins que des réactions provenant du parler d'autres localités n'interviennent, le parler de chaque localité tend donc à prendre progressivement des caractères propres.

Sans doute ces innovations locales, procédant de conditions inhérentes à la langue ou à un ensemble de populations, peuvent être et sont, en effet, généralement communes à tout un groupe de localités. Mais, de ce que une ou plusieurs innovations identiques ont lieu indépendamment dans certaines localités, il ne résulte pas que telle autre innovation réalisée dans l'une de ces localités ou dans quelques-unes d'entre elles doive se retrouver dans toutes les autres. Chaque innovation a ses limites géographiques propres. Ceci se traduit de la manière suivante : si l'on

examine à une date donnée tous les parlers d'une région où a dû être parlée à une date antérieure une même langue et où, dans la mesure du possible, chaque localité a eu un développement linguistique autonome (notion toute relative, bien entendu), on observe qu'à une forme de la première époque répondent des formes diverses suivant les localités, et que chacune des particularités divergentes a son extension propre. Soit par exemple la région du département français des Landes étudiée par M. Millardet, dans son *Petit atlas linguistique des Landes* ; soit le mot latin vulgaire *yugu* « joug » ; la région étudiée se divise en quatre parties, disposées de la manière suivante (réduite à un pur schéma), en tenant compte des formes prises par le mot :

La ligne qui marque les limites des traitements *j* et *y* du *y* initial de *yugu* coupe celle qui marque la limite des traitements *iw* et *u* de la suite du mot (p. 245 du livre cité). Les exemples de ce genre abondent, et le livre de M Millardet en fournit à lui seul un bon nombre. Il résulte de là que même des localités qui présentent en commun plusieurs innovations linguistiques ont cependant, après quelques siècles de développement indépendant, des parlers distincts.

Ce type de différenciation linguistique atteint surtout la prononciation et la grammaire ; il touche relativement peu le vocabulaire. Il résulte de la suite naturelle des générations et se réalise par le fait que la transmission du langage est discontinue : la continuité des langues n'est en effet assurée que par une suite de créations successives, dont chacune apporte inévitablement quelques différences avec l'état antérieur.

Un tout autre type de différenciation a pour cause l'existence

de distinctions à l'intérieur d'un groupe social étendu. Aucune population n'est tout à fait homogène ; chaque différenciation sociale a chance de se traduire par une différenciation linguistique. D'une part, il y a des groupes durables. Ainsi les femmes et les hommes forment, surtout dans les sociétés de civilisation inférieure, deux groupements fortement distincts, et qui ont parfois des langues différentes, cas observé surtout dans certaines régions de l'Amérique du Sud ; sans aller à cette extrémité, les hommes et les femmes ont des différences de parler qui sont parfois sensibles, même chez les peuples européens. Toute société tend aussi à constituer des classes distinctes, et au fur et à mesure que les membres de chacune de ces classes tendent à vivre entre eux, en se séparant des autres classes, ils se donnent des parlers différents ; les diverses classes sociales d'une grande ville moderne tendent à occuper des quartiers différents ; de plus en plus, elles n'ont entre elles que des rapports extérieurs ; il y a souvent assez loin de la langue de la bourgeoisie dans une grande ville moderne à la langue des ouvriers ; et ni la bourgeoisie, ni les ouvriers ne forment des unités ; il y a des classes et des sous-classes, chacune avec ses particularités linguistiques. On a observé de même des particularités de langue dans les castes hindoues ; les anciens dramaturges de l'Inde prêtaient des parlers différents aux gens des diverses castes, depuis les brahmanes et les rois, qui parlent sanskrit, jusqu'aux gens de castes inférieures, qui se servent de prâkrits fortement dégradés. Là où existe une division du travail, chaque profession a nécessairement ses termes propres qui ne sont pas courants, ou même sont inconnus, en dehors de la profession ; cette différence se combine avec celle du sexe ; car, chez les peuples peu civilisés, les hommes et les femmes remplissent d'ordinaire des fonctions distinctes, les hommes chassant et s'occupant du bétail par exemple, et les femmes recueillant ou cultivant des végétaux. Plus la société se complique, plus les métiers deviennent nombreux et distincts les uns des autres, plus ils absorbent entièrement l'activité des individus qui les pratiquent, et plus par suite se spécialisent aussi les langues techniques correspondantes. On a peine à imaginer seulement la variété

des langues spéciales employées chez un grand peuple moderne, et la lecture des manuels techniques laisse deviner l'existence de vocabulaires très distincts dont personne ne possède ni la totalité ni même une partie notable. Les malfaiteurs, les vagabonds, les mendiants ont partout des argots, caractérisés par des dénominations spéciales des objets qui les intéressent particulièrement.

Il n'y a pas à tenir compte seulement des professions qui prennent la part principale de l'activité de chaque homme. Il y a quantité d'occupations et d'activités transitoires qui suffisent à déterminer des parlers spéciaux. Les cérémonies religieuses appellent presque toujours une langue différente de celle que l'on emploie dans l'usage ordinaire de la vie : destinées à faire pénétrer l'homme dans un monde séparé, celui du sacré, elles exigent une langue aussi séparée, une langue sacrée ; et, un peu partout, on constate que, lorsqu'ils accomplissent des rites, les hommes recourent à des manières de parler spéciales. Ceci n'est pas seulement le fait des non civilisés, et la civilisation tend même à exagérer cette tendance : les églises chrétiennes ont chacune une langue religieuse, qui est au fond la même pour tous les chrétiens appartenant à une même confession et qui n'a rien de commun avec le langage courant des fidèles : l'église romaine célèbre partout ses offices en latin (sauf dans quelques groupes orientaux où est concédé l'usage de langues liturgiques, dont aucune du reste n'est identique à celles que parlent aujourd'hui les fidèles) ; l'église grecque a gardé le grec ancien ; les églises slaves se servent non du russe, du serbe, etc., mais du slavon ecclésiastique ; etc. L'arabe littéral est partout la langue de religion de l'Islam ; le pâli, la langue du bouddhisme méridional. C'est l'état normal là où existe un clergé spécialisé dont la seule fonction est la fonction religieuse et qui tend à augmenter de plus en plus tout ce qui le distingue du commun du peuple. Mais, chez les non civilisés, la religion est mêlée à la vie entière ; la langue religieuse n'est pas aussi entièrement spécialisée ; toutefois, elle tend à être distincte, et, comme la religion intervient constamment, la langue religieuse est souvent employée. Du reste, ce ne sont pas seulement les cérémonies religieuses qui

appellent une langue spéciale ; d'autres actes collectifs comportent aussi une manière particulière de parler ; ainsi pour la chasse, pour certaines récoltes, on recourt à des vocabulaires spéciaux ; ou bien des événements qui exigent des pratiques compliquées de purification entraînent en même temps des interdictions de vocabulaire temporaires ou durables qui s'étendent à une famille, à un village, à une tribu ; il suffit d'un mot interdit à certains individus pour occasionner la formation de mots nouveaux et des changements de sens. Chez les civilisés, des groupements transitoires produisent des effets analogues ; les jeunes gens groupés à l'école, au régiment, etc. se créent peu à peu des manières de parler qui leur sont propres. Les associations, les groupes sportifs s'expriment aussi d'une façon spéciale, au moins dans tous les moments où le groupe est réuni en totalité ou en partie.

On aura une idée de l'importance des langues spéciales si l'on note qu'une population du Sud de l'Inde, celle des Todas qui ne compte qu'environ 800 individus, a trois langues religieuses spéciales, un argot et une division dialectale d'origine sociale [1].

Sauf les cas où il s'agit de langues entièrement distinctes, c'est-à-dire sauf le cas de certaines langues religieuses, ces différenciations d'origine sociale atteignent avant tout le vocabulaire ; elles intéressent beaucoup moins la prononciation et la grammaire. Les langues spéciales, celles des métiers, celles des groupements transitoires, celles même de la religion, ne se distinguent le plus souvent que par des vocables particuliers ; le système général de la langue, caractérisé par la prononciation et par les formes grammaticales, reste un. Ce type de différenciation s'oppose donc à celui qui résulte de l'emploi qui est fait du langage et de la suite des générations ; il se présente souvent à l'intérieur d'une même localité, et il atteint des éléments linguistiques tout autres.

Quelles qu'elles soient, les différenciations tendent à rendre la

1. D'après RIVERS, *The Todas*.

langue intelligible seulement à des groupes sociaux de plus en plus étroits. Elles vont ainsi contre l'objet du langage qui est de faciliter les relations entre les hommes. Quand il ne se produit pas de réactions, la différenciation aboutit à des résultats tels que l'utilité du langage en est singulièrement diminuée.

Les populations qui vivent isolées, divisées en petits groupes ayant peu de rapports entre eux, peuvent arriver à posséder des parlers qui ne sont pas compris hors de tribus parfois très peu nombreuses. Ainsi les indigènes de l'Amérique présentent une variété infinie de langues dont beaucoup ont quelques traits communs, mais qui diffèrent assez dans le détail, non seulement pour qu'on ne se comprenne pas de tribu à tribu, mais même pour que, au moins à première vue, les linguistes soient hors d'état d'établir un classement. Il y a en Amérique des centaines de langues diverses, qui forment un grand nombre de groupes non encore rapprochés les uns des autres, simplement parce que la population indigène semble toujours avoir été peu dense dans la plus grande partie du continent américain, que les tribus qui parlent ces langues sont demeurées isolées les unes des autres et que toutes les forces de différenciation ont agi sans se heurter à une résistance efficace.

En Europe, la dissolution de l'empire romain a permis la différenciation du latin parlé dans l'Ouest de l'Empire en un grand nombre de parlers distincts ; au IV^e^ siècle ap. J.-C., il n'y avait encore qu'un latin, les différences locales étant minimes et ne gênant nulle part la compréhension ; au X^e^ siècle, les langues romanes avaient déjà reçu leurs traits essentiels, et le parler d'un Parisien était devenu inintelligible à un Romain. Les parlers locaux ont depuis continué d'évoluer d'une manière autonome, et la différenciation est parvenue à ce point qu'il existe aujourd'hui en France une multitude de parlers locaux qui tous sont du latin transformé et qui néanmoins sont trop distincts pour permettre à ceux qui les parlent de s'entendre entre eux. Ceci tient à ce que, durant longtemps, les rapports entre les hommes qui employaient ces parlers ont été peu fréquents et peu importants, et à ce que, de bonne heure, il s'est formé des langues

communes : latin médiéval et français commun, qui ont servi de moyens de communication à tous les individus qui avaient entre eux des rapports suivis. Réservés à l'usage local, les parlers locaux ont pris de plus en plus un caractère particulier, et l'importance même du rôle pris par la langue commune a eu pour conséquence une résistance moindre du parler local à l'isolement. On tend vers une situation où chaque petit groupe humain, composé souvent de quelques individus seulement, aurait un idiome qui lui serait entièrement propre.

Pareille situation est instable, et il y est mis fin de deux manières : par extension d'une langue nouvelle et par substitution d'une langue commune à des parlers de même famille.

D'une part, des populations fragmentées en petits groupes distincts et parlant des langues distinctes sont hors d'état de résister à l'attaque de groupes unis et comprenant des individus nombreux bien organisés ; ces populations sont donc conquises, et elles sont ou détruites ou assimilées : les indigènes de l'Amérique ont été refoulés par les Européens ; les uns ont été tués, d'autres dépouillés de leurs moyens d'existence, d'autres assimilés ; plusieurs de leurs langues sont déjà disparues ; d'autres ne sont plus parlées que par quelques individus ; et trois langues, l'anglais, l'espagnol et le portugais (pour ne rien dire du français canadien, trop peu important et qui ne saurait se maintenir définitivement), se partagent l'Amérique entière ; les anciennes langues n'y ont plus qu'une place insignifiante à tous égards. Des faits pareils se sont produits à toutes les périodes de l'histoire : l'Italie avait, au IIIe siècle av. J.-C., quantité de langues distinctes : le latin n'y occupait qu'une place restreinte autour de Rome, et encore le latin de Préneste, par exemple, différait-il beaucoup de celui de Rome : ce qui était à Rome *lûna* « lune » se prononçait *lôsna* à Préneste. D'autres langues encore plus différentes du latin, mais cependant apparentées, se parlaient au Nord-Est et au Sud : l'ombrien et l'osque forment un groupe où l'on distingue du reste plusieurs parlers différents, bien qu'on n'ait de renseignements que sur un petit nombre de localités. En outre, on employait au Sud

des parlers grecs les uns du type dorien, les autres du type de la *koinè* ionienne-attique. L'étrusque existait encore au nord de Rome, et plus au Nord le gaulois. D'autres langues plus obscures, notamment le messapien et le vénète, étaient en usage au Sud-Est et au Nord-Est ; le ligure subsistait aussi le long du golfe de Gênes ; enfin on a trouvé tout au Nord quelques inscriptions en une langue mal déterminée qui n'est peut-être aucune de celles qui viennent d'être nommées. La généralisation de l'emploi du latin a mis fin à ce morcellement linguistique.

La conquête arabe, en substituant l'arabe à l'araméen en Syrie, au copte en Égypte, au berbère dans l'Afrique mineure et à nombre de petits idiomes, a créé une vaste unité ; le copte est sorti de l'usage parlé ; l'araméen et le berbère n'existent plus qu'à l'état de parlers locaux et s'éliminent peu à peu ; déjà il ne subsiste presque plus rien de l'araméen. Or, l'araméen s'était lui-même substitué antérieurement à une série d'autres langues, notamment à l'hébréo-phénicien et à l'assyro-babylonien.

Pour que des généralisations de ce genre aient lieu, il n'est ni nécessaire ni suffisant qu'il y ait conquête. L'araméen s'est étendu sans conquête, sans domination politique, simplement parce qu'il était l'idiome de l'administration et des affaires. En revanche, si le latin est devenu la langue de tout l'Occident de l'empire romain, il n'a pu se répandre dans les parties orientales de l'empire, bien que la puissance romaine n'y ait pas été moindre qu'en Occident. Pour qu'une langue se généralise, il faut et il suffit qu'elle serve de support à une civilisation. Le latin a pu se répandre partout où il servait à porter la civilisation gréco-romaine ; il n'a guère pénétré là où cette civilisation existait déjà, sous la forme hellénique. L'arabe a été généralisé par la conquête ; mais cette conquête comportait une forme spéciale de civilisation, fondée sur les civilisations araméenne, iranienne et byzantine ; et c'est cette civilisation qui a donné à la langue arabe sa puissance d'expansion ; on le voit aujourd'hui encore, maintenant que l'arabe, malgré la déchéance politique, refoule peu à peu le berbère en Algérie, dans une colonie française, où le français est la langue de l'administration, des chemins de fer et de

toutes les affaires importantes. Le russe [1] a une force d'expansion considérable au Caucase, où il est le porteur de la civilisation occidentale ; il n'en a presque aucune dans un vieux pays de civilisation occidentale comme la Pologne. Il est permis de se demander en quelle mesure, malgré l'unification politique et malgré l'école, le turc pourra devenir la langue commune de l'empire turc et déplacer l'arabe, l'albanais, le grec, l'arménien, le judéo-espagnol et les autres idiomes parlés dans l'empire ; le turc n'a pas sur ces langues l'avantage d'apporter une forme supérieure de civilisation ; et pourtant la situation linguistique de l'empire turc est intolérable, et ne saurait persister, maintenant que les communications deviennent courantes et les affaires communes nombreuses.

Les conditions politiques, économiques, religieuses qui déterminent la généralisation d'une langue sont complexes et diverses ; on n'a jamais eu occasion de les examiner en détail, car, pour les expansions qui ont eu lieu dans le passé, on n'a que des données insuffisantes sur le procès de l'extension ; et, pour celles qui ont lieu actuellement, on ne les a pas étudiées en détail. L'un des pays où il serait le plus facile de les observer est la Russie ; car le russe est sans doute l'une des langues qui gagnent le plus à l'époque moderne ; il élimine peu à peu les parlers finnois dans toute la partie de la Russie où ils subsistent encore (abstraction faite, bien entendu, du finnois proprement dit au Nord, dans la Finlande, qui est demeurée un pays autonome jusqu'à présent) ; il devient au Caucase la langue commune ; il progresse dans l'Asie centrale et surtout en Sibérie. Mais personne n'a étudié de près la marche de ces phénomènes. Les faits sont du reste différents suivant les régions : en Sibérie, il s'agit surtout de colonisation, et le russe est porté par des paysans. Au Caucase, au contraire, il s'agit de civilisation ; c'est par l'école, par le gymnase, par l'Université, et aussi par l'armée, par l'administration, par les chemins de fer, par les exigences des affaires, qui se font en russe, que le russe est porté ; la bourgeoisie arménienne ne parle

1. Cet article a été écrit en 1911.

plus que russe, alors que la population rurale arménienne ne parle encore que l'arménien ; dans une grande ville géorgienne, à Tiflis chez les Arméniens (qui forment une colonie très nombreuse et jouant un grand rôle), la population ouvrière et les petits boutiquiers de langue arménienne s'opposent à la bourgeoisie de langue russe, et dont beaucoup de membres ne comprennent même pas l'arménien.

Dans tous les cas, il y a un trait commun : la puissance d'une organisation politique et la valeur d'une civilisation peuvent en être les causes prochaines ; mais la cause profonde qui détermine le phénomène est l'utilité singulière que présente une langue employée sur un vaste domaine. Plus vaste est ce domaine, plus importantes sont les relations soutenues par les hommes qui l'habitent, et plus le besoin d'une langue commune se fait sentir, plus l'expansion de la langue du pouvoir ou de la civilisation dominante est facile et rapide.

La puissance des mouvements de ce genre n'est pas liée, comme on pourrait le croire, à l'emploi fréquent de l'écriture ou à l'existence d'écoles. C'est sans écriture, et par suite sans écoles, que l'indo-européen, sous les formes diverses qu'il a prises, a couvert l'Europe entière, que le bantou a conquis presque toute l'Afrique méridionale, que le berbère s'est étendu sur toute la largeur de l'Afrique du Nord. Partout où l'on observe une même langue peu différenciée sur un domaine étendu, il y a lieu de supposer que l'expansion est de date relativement récente ; car la différenciation procède en général assez vite, et là où l'on peut suivre les faits de près, il suffit de quelques siècles pour transformer profondément une langue, au moins dans les pays où il y a eu mélange de populations autrefois différentes, c'est-à-dire presque partout. Les groupes où l'on observe des dialectes très pareils les uns aux autres résultent donc pour la plupart d'extensions peu anciennes, et dont on n'ignore la date que parce que l'histoire commence très tard pour la plupart des peuples et des langues.

L'extension d'une langue entièrement nouvelle n'est que l'un des moyens par lesquels se réalise l'unité de langue sur un domaine étendu. Pour se produire, la réaction contre la différenciation

n'attend souvent pas que les parlers soient devenus bien distincts et que les sujets parlants cessent tout à fait de s'entendre entre eux ; cette réaction est, dans bien des cas, constante tout comme la tendance à la différenciation. Une fois une unification réalisée, les effets ont chance de persister aussi longtemps que persistent les causes de l'unification. Le latin est demeuré, avec des changements minimes et pour la plupart communs à tout le domaine, la langue de l'empire romain occidental du Ier siècle av. J.-C. au Ve siècle après. Le français, fixé du XIVe au XVIIe siècle, ne se modifie que lentement, et les modifications s'étendent en général à tous les individus qui l'emploient. Dans de grandes langues communes de civilisation, la résistance à l'innovation est forte, parce que l'innovation doit s'étendre à un grand nombre de sujets répandus sur une aire géographique très vaste. Et c'est un grand bien. Il est permis de souhaiter que l'anglais ne vienne pas à se différencier trop fortement en Amérique ou en Australie, et que les divergences déjà sensibles que l'indépendance politique, la différence de situation et les origines variées des sujets parlants ont introduites entre le castillan et la langue du Chili ou de la République Argentine ne s'accentuent pas de manière à produire des idiomes nouveaux. C'est le rôle de l'école et de la littérature de maintenir les unités linguistiques une fois créées.

Dans les domaines où les parlers locaux appartiennent à une même langue ancienne, l'unité brisée se laisse assez aisément rétablir. Soit une région où s'emploient des parlers qui sont des formes distinctes qu'a prises une même langue par le fait d'évolutions indépendantes, la France du Nord par exemple ; si l'on considère les extrémités du domaine, on rencontre des parlers très différents, et un Franc-Comtois ou un Lorrain ne comprend pas un Picard qui ne comprend pas un Berrichon ; mais il y a d'un groupe à l'autre une série de transitions ; on ne peut tracer nulle part de limites précises entre dialectes, et les habitants d'un village comprennent toujours ceux d'un village voisin, chacun employant son propre parler. Il y a alors des règles de correspondance dont les sujets parlants ont conscience, et qui leur donnent le moyen de transposer en gros un parler dans l'autre. Il peut

se constituer une langue commune en opposition avec tous les parlers locaux : ce n'est pas un idiome nouveau, puisque l'on passe de cette langue commune au parler local, et inversement, au moyen de règles de transposition. Ce phénomène a pris dans l'Europe actuelle une importance décisive ; il en domine présentement tout le développement linguistique ; mais, à des degrés variés, il est de tous les temps. Dès l'instant que des hommes appartenant à des groupes divers emploient des parlers déjà différenciés, ils ont le sentiment de ces règles de correspondance : un Ionien savait qu'à ses *e* correspondent, dans des cas définis, des *ā* doriens ou éoliens. Et il est inévitable que, parmi les parlers en usage, il y en ait qui appartiennent à des groupes plus puissants ou supérieurs en civilisation, doués d'un prestige supérieur pour quelque raison que ce soit. Ces parlers servent de modèle aux autres ; on vise à s'en rapprocher, sinon à les parler exactement, dans les relations entre groupes. C'est le commencement de l'évolution qui conduit à créer une langue commune sur la base de l'un des parlers du groupe et à éliminer tout ou partie des innovations étroitement locales.

L'histoire de ces créations de langues communes est diverse et souvent très compliquée.

Le type le plus simple est celui du français : le parler d'une région centrale, qui est celui des chefs du pays et où la civilisation a son centre, devient intégralement la langue commune ; Paris, résidence principale du roi de France, centre naturel de la France du Nord, siège d'une Université puissante qui a eu au moyen âge une forte influence, a donné son parler à la royauté française ; dès le moyen âge, le français commun est la langue de Paris ; les textes écrits en d'autres dialectes n'ont qu'une importance secondaire et, de bonne heure, le français écrit n'est rien que la langue de Paris, telle qu'elle se fixe sous toutes sortes d'influences savantes et littéraires et telle que l'adopte l'administration royale. Les parlers du Midi de la France appartenant à des types tout autres, inintelligibles aux Français du Nord, n'ont exercé aucune action : dans toute la France méridionale, le français est une langue étrangère qui s'impose aux villes, mais qui n'a pas encore déplacé les parlers locaux à la campagne.

La situation de l'anglais est un peu différente; l'anglais commun est en somme le parler de Londres, comme le français est le parler de Paris; mais, à la différence de Paris, Londres se trouve au point de rencontre de plusieurs dialectes, si bien que l'anglais présente des traits qui appartiennent originairement à des dialectes distincts.

Le cas de l'allemand est tout autre : il n'existait en Allemagne aucun centre, comparable à Paris ou à Londres; aucun ne s'imposait naturellement. Mais le besoin était le même. La langue commune est sortie du grand mouvement de colonisation qui a permis aux Allemands de conquérir progressivement toute l'Allemagne orientale et qui est le fait fondamental de l'histoire de l'Allemagne; c'est en réalisant cette œuvre que les Allemands formaient une unité; et c'est par suite dans les pays de colonisation que l'allemand littéraire s'est formé, surtout parmi la bourgeoisie des villes établies dans les régions colonisées au moyen âge; c'est donc en Bohême et surtout dans les villes de Saxe que l'allemand actuel s'est constitué au cours du moyen âge. Les chancelleries des princes et des villes l'ont fixé peu à peu. La Réforme, en le prenant pour langue littéraire, lui a donné sa forme définitive. Mais l'allemand commun est demeuré, plus que le français et l'anglais, une simple langue écrite; les choses qui se fixent par l'écriture, à savoir le vocabulaire et la grammaire, sont à peu près identiques partout; mais la prononciation varie beaucoup de région à région, sans qu'on puisse nettement, comme en français ou en anglais, déclarer vicieuse telle ou telle manière de prononcer : tout le monde sait combien le système articulatoire est différent à Cologne, à Berlin, à Leipzig, à Francfort, à Munich et à Vienne.

Les voies par lesquelles se réalise une langue commune superposée aux parlers locaux sont multiples et embrouillées. L'un des cas les mieux étudiés et les plus curieux est celui du grec ancien. A la date où apparaissent les premiers textes grecs, vers le VII[e] siècle av. J.-C., la Grèce est divisée en cités qui constituent autant de petits États autonomes, dont quelques-uns seulement tendent à se réunir en confédérations. A en juger par les textes

du v^e^ siècle av. J.-C., il y avait à ce moment autant de parlers que de cités, surtout chez les populations doriennes. Toutefois, là où la civilisation avait commencé à se développer largement, en Asie Mineure, il s'était constitué des langues communes : toute l'île de Lesbos paraît n'avoir eu qu'une seule langue écrite, celle que l'on connaît par les poètes Alcée et Sapho ; l'épopée et la poésie didactique se servent d'une langue spéciale, mélange singulier d'éolien et d'ionien, langue artificielle, mais qui sert à tous les Grecs, même hors d'Asie Mineure, comme on le voit par Hésiode ; les cités ioniennes d'Asie emploient toutes, dans leurs inscriptions, un même ionien commun, et c'est cette langue qui, à quelques détails près, sert aussi à la prose d'Hérodote et du dorien Hippocrate, à la poésie d'Archiloque et d'Anacréon. Les Doriens, relativement barbares, ont encore autant de langues officielles distinctes que de cités. La fondation de l'empire athénien et le puissant développement de la civilisation d'Athènes amènent ensuite la création de la prose attique, qui se substitue dans l'usage à la prose ionienne ; la ruine de l'indépendance de l'Ionie et l'arrêt du développement de sa civilisation ôtaient à l'ionien tout prestige ; quand la Macédoine se civilise, la cour prend pour langue l'attique, si bien qu'on ne possède pas une ligne de texte en macédonien et qu'on ne sait même pas à quel groupe de langues indo-européennes appartenait l'idiome propre des Macédoniens. Athènes pouvait dès lors perdre son influence politique ; les conquêtes d'Alexandre et la fondation de royaumes coloniaux ont eu pour conséquence l'extension lointaine de l'attique un peu mêlé d'ionien qu'on connaît sous le nom de langue commune hellénique, de *koinê*. Incidemment il s'était développé, dans les régions où dominait le dialecte dorien ou les parlers très pareils du Nord-Ouest, une autre langue commune, de type différent ; en Sicile et dans les pays où ont dominé aux III^e^ et II^e^ siècles av. J.-C. les confédérations étolienne et achéenne, il y a donc eu des langues communes différentes de la *koinê* ionienne-attique. Mais ces langues n'ont eu que très peu accès à la littérature. En ruinant les confédérations, la conquête romaine a provoqué l'anéantissement de leurs langues. Et il n'a subsisté que

la grande langue de civilisation, la *koiné* proprement dite, qui perdait de plus en plus ses traits proprement athéniens et devenait l'idiome commun à toute la partie orientale de l'empire romain. Les cités, qui avaient cessé d'être des États indépendants, acceptent progressivement la langue commune qui se répand même à la campagne, et les parlers grecs locaux disparaissent définitivement dans les premiers siècles de l'ère chrétienne. L'existence d'un grec commun ne devait pas du reste être de longue durée; car la dissolution progressive de l'empire, en rompant peu à peu les relations entre les provinces, recréait des conditions favorables à une différenciation; et ainsi se sont développés, dès le VIII^e^-IX^e^ siècle ap. J.-C. au moins, et sans doute avant, les parlers grecs modernes, de nouveau distincts les uns des autres, mais auxquels on s'efforce de superposer maintenant une nouvelle langue commune, fondée en grande partie sur la tradition littéraire [1].

L'action d'une langue commune ainsi superposée aux parlers locaux tend à supprimer les différences locales, et elle y arrive souvent d'une manière complète, comme dans le cas du grec ancien qui vient d'être cité. Mais cette suppression n'a lieu que d'une manière progressive; les mots, les prononciations, les formes grammaticales de la langue commune remplacent peu à peu dans chaque localité les formes indigènes. Les textes grecs du V^e^ siècle av. J.-C. sont vraiment locaux; mais, à partir du moment où l'alphabet ionien se généralise et où l'on abandonne les alphabets locaux, où se marque ainsi la tendance vers une civilisation hellénique commune, on voit les formes ioniennes-attiques pénétrer partout peu à peu, et les traits propres de chaque parler s'effacent les uns après les autres. Là même où les patois ont, dans la France du Nord, l'air de subsister, ils sont en réalité tout pénétrés de français commun. Quand un patois existe concurremment avec une langue commune qui gagne progressivement, tout le prestige est pour la langue commune dont les

1. Sur tous ces faits, voir A. Meillet, *Aperçu d'une histoire de la langue grecque*, 2^e^ édition, 1920.

mots entrent dans le parler local. Parler patois, c'est alors souvent patoiser la langue commune, et il arrive qu'on la patoise à faux. On a nombre de cas où des Grecs non ioniens ont, pour parler dialecte, mis à faux des *á* au lieu de l'*é* ionien parce qu'ils savaient que l'ionien a *é* au lieu de la plupart des *â* dialectaux. Avant de disparaître, les patois se dépouillent ainsi peu à peu de leur individualité. On se rend compte de plus en plus, par la comparaison des aires géographiques qu'occupe chaque mot, que le vocabulaire des patois français se compose en grande partie de mots qui se sont répandus à des dates diverses bien postérieures au v^e^ siècle ap. J.-C. et que le nombre des vocables d'un patois qui remontent directement au latin, sans aucun emprunt, est petit. On a pu dire qu'on n'est sûr, presque pour aucun mot d'un parler local français, qu'il résulte de la transmission ininterrompue de génération en génération depuis le latin vulgaire jusqu'à la forme patoise actuelle.

En effet, ce n'est pas seulement la langue commune qui agit sur le parler local. C'est souvent d'abord la langue d'un parler local plus important. Il y a ainsi des croisements presque inextricables d'influences successives ou simultanées. L'hypothèse du développement autonome qu'on a envisagée ci-dessus est presque toujours une simple fiction, commode pour l'étude, et qui fournit une première approximation des faits, mais à laquelle on doit renoncer quand on veut serrer les données de près.

Par suite des groupements partiels qui résultent de ces actions de centres secondaires, ou de l'absence d'un groupe principal et de l'existence d'un certain nombre de centres indépendants, les parlers appartenant à une population où les rapports sont courants, et qui forme en quelque mesure une unité, tendent à présenter des traits communs; ils forment un *dialecte*. Les limites des développements autonomes envisagés plus haut sont indépendantes pour chacun des faits particuliers, on l'a vu; au contraire les limites de faits qui résultent des unifications considérées ici tendent à coïncider les unes avec les autres. Et il se constitue des dialectes bien définis, d'autant plus définis que la vie provinciale a plus de réalité. Le centre de la France, où Paris est la

seule ville dominante, n'a pas de dialectes; le Midi, où il y a des provinces caractérisées, a au contraire des dialectes bien caractérisés aussi : provençal, gascon, dont les limites se laissent tracer avec une certaine netteté. Et l'on retrouve de véritables dialectes en Normandie, en Picardie, en Lorraine, en Franche-Comté.

Totale ou partielle, l'unité linguistique qui se constitue par ces actions diverses n'exprime pas une unité d'origine des populations qui parlent la langue désormais unifiée; elle traduit l'existence de rapports sociaux importants ayant existé à un certain moment, le sentiment d'une unité de civilisation. Les unités linguistiques ne perdent rien de leur intérêt à être considérées de cette manière; elles servent à indiquer des mouvements de civilisation, souvent très complexes, plutôt que des communautés d'origine.

La tendance à l'unité de langue là où il y a unité de civilisation est si forte qu'une certaine espèce d'unité tend à se réaliser, même à travers des idiomes profondément distincts et qui restent distincts. Les grandes langues communes de l'Europe actuelle forment à tous égards des systèmes absolument différents; elles ont des prononciations et des grammaires strictement autonomes. Mais ces langues reposent toutes sur un même fonds de civilisation, et il est aisé de constater qu'elles présentent en grande quantité des éléments communs. D'abord, par emprunt des unes aux autres, ou par suite de leur unité d'origine indo-européenne, elles ont en commun beaucoup de mots; quand, pour constituer des langues artificielles, on a dressé le bilan des mots communs à l'italien, à l'espagnol, au français, à l'anglais, à l'allemand et au russe, on a trouvé assez de termes communs à quatre ou cinq de ces langues pour constituer un vocabulaire où, par suite des emprunts innombrables de l'anglais et des emprunts assez nombreux de l'allemand au latin et aux langues néo-latines, le latin est l'élément essentiel, et où le russe, demeuré longtemps en dehors du grand courant de la civilisation européenne, ne fournit rien. En second lieu, les manières de parler ont été traduites d'une langue dans l'autre; le grec *syneidêsis*, le latin *con-*

scientia (français *conscience*), l'allemand *gewissen*, le polonais *sumienie*, le russe *sovêst'* sont autant de mots distincts au premier abord ; mais il suffit de les analyser pour apercevoir qu'ils se superposent exactement et présentent un même mode de formation, résultant de ce qu'ils ont été calqués les uns sur les autres. Des mots allemands comme *Ausdruck* et *Eindruck* calquent les mots *expressio* et *impressio* courants en latin savant, et il est curieux de voir comment depuis le XVIII^e^ siècle, les sens de *Ausdruck* reproduisent toutes les nuances de sens du fr. *expression*. On peut donc dire qu'il s'est produit par-là, en un certain sens, une unification des langues européennes ; cette unification a commencé lorsque s'est fondée la civilisation méditerranéenne, avant même l'arrivée des Hellènes en Grèce ; elle s'est continuée par l'hellénisation du latin et n'a jamais cessé depuis.

Ce sont ces divers faits d'unification qui rendent possible la linguistique historique. Sans les unifications successives qui ont maintenu ou établi constamment l'unité de langue sur de vastes domaines, les linguistes se trouveraient devant une poussière de parlers avec laquelle il serait impossible d'opérer. L'extension de l'indo-européen à une partie de l'Asie et à presque toute l'Europe fournit la base de la grammaire comparée des langues indo-européennes : une première différenciation a créé des parlers qui se sont à leur tour unifiés en groupes nouveaux : indo-iranien, slave, germanique, hellénique, italique, celtique, etc. ; chacune des langues communes qui se sont constituées on ne sait comment s'est à son tour imposée à un domaine étendu, puis s'est différenciée à son tour. Entre l'indo-européen et le français moderne, on entrevoit ainsi toute une série d'unifications et de différenciations successives : une unité italo-celtique, qui se brise et aboutit à la création d'une unité italique et d'une unité celtique ; une unité latine provenant d'une différenciation de l'unité italique ; une unité latine, brisée en parlers infiniment divers ; de là est issu, entre autres langues, le français dont l'unification se poursuit encore.

On voit comment, si le français est une forme prise par l'indo-européen, les hommes qui parlent français aujourd'hui n'ont pas

reçu leur langue de leurs ancêtres par une transmission ininterrompue. Il y a eu constamment emprunt de langues communes. Le progrès de la linguistique tend à mettre de plus en plus en évidence ce fait dominant de l'histoire des langues : la création et l'extension de ces langues communes, qui sont le produit de l'unité de civilisation, sur des domaines plus ou moins vastes. Ainsi apparaît le caractère éminemment social du développement des langues. Les innovations linguistiques procèdent, en partie, de faits anatomo-physiologiques et psychiques ; mais ce qui fixe les formes et détermine le développement, ce sont les conditions sociales où se trouvent les sujets parlants.

Après avoir longtemps cherché de tous côtés le développement naturel du langage, les linguistes ont fini par reconnaître qu'on ne l'observe exactement nulle part et que toutes les langues connues, populaires ou savantes, trahissent la préoccupation d'un mieux dire qui partout a conduit les sujets parlants à emprunter le langage de ceux qui sont censés parler mieux. Chaque différenciation est tôt ou tard, et parfois immédiatement, suivie d'une réaction qui tend à rétablir ou à instaurer l'unité de langue là où il y a unité de civilisation.

Linguistique historique et linguistique générale.

L'ÉVOLUTION

DES

FORMES GRAMMATICALES [1]

Les procédés par lesquels se constituent les formes grammaticales sont au nombre de deux ; tous les deux sont connus, même des personnes qui n'ont jamais étudié la linguistique, et chacun a eu occasion, sinon d'y arrêter son esprit, du moins de les observer en passant.

L'un de ces procédés est l'analogie ; il consiste à faire une forme sur le modèle d'une autre ; soit par exemple les types français : *nous finissons, vous finissez, ils finissent* ; *nous rendons, vous rendez, ils rendent* ; *nous lisons, vous lisez, ils lisent* ; sur *nous disons, ils disent*, l'enfant qui apprend à parler est conduit à former *vous disez* sans avoir jamais entendu pareille forme : c'est une forme dite *analogique*. Toutes les formes régulières de la langue peuvent être qualifiées d'analogiques ; car elles sont faites sur des modèles existants, et c'est en vertu du système grammatical de la langue qu'elles sont recréées, chaque fois qu'on en a besoin. Mais d'ordinaire ces formes régulières sont aussi celles qu'on a eu occasion d'observer, et, sauf quand il s'agit de mots nouveaux ou rares, la forme obtenue par le fonctionnement du système grammatical reproduit le plus souvent une forme déjà entendue et enregistrée dans la mémoire. La tradition est d'accord avec les exigences du système. Mais il arrive, comme dans le cas cité, que la tradition et le système ne soient pas d'accord, et que, étant donné l'état de la langue à un moment donné, il y ait plusieurs formes possibles ; alors l'analogie produit des formes nouvelles, indépendantes de la tradition. Et c'est dans

1. *Scientia* (*Rivista di scienza*), vol. XII (1912), nº XXVI, 6.

ces cas qu'on parle d'ordinaire de *formes analogiques*; il serait plus juste de dire : *innovations analogiques*.

L'autre procédé consiste dans le passage d'un mot autonome au rôle d'élément grammatical. Par exemple *suis* est un mot autonome dans la phrase, du reste très artificielle, *je suis celui qui suis*, et a encore une certaine autonomie dans une phrase telle que : *je suis chez moi*; mais il n'est presque plus qu'un élément grammatical dans : *je suis malade, je suis maudit*, et il n'est tout à fait qu'un élément grammatical dans : *je suis parti, je suis allé, je me suis promené*, où personne ne pense ni ne peut penser à la valeur propre de *suis*, et où ce que l'on appelle improprement l'auxiliaire n'est qu'une partie d'une forme grammaticale complexe exprimant le passé. Il est pourtant clair — et l'histoire de la langue montre de manière évidente — que *suis* est dans : *je me suis promené* le même mot que dans : *je suis ici*; mais il est devenu une partie constituante d'une forme grammaticale.

Ces deux procédés, l'innovation analogique et l'attribution du caractère grammatical à un mot jadis autonome, sont les seuls par lesquels se constituent des formes grammaticales nouvelles. Les faits de détail peuvent être compliqués dans chaque cas particulier; mais les principes sont toujours les mêmes.

Depuis qu'il existe une grammaire comparée systématique, on a été amené à attribuer plus d'importance tantôt à l'un et tantôt à l'autre des deux procédés. Le fondateur de la grammaire comparée des langues indo-européennes, Fr. Bopp, croyait que l'examen des plus anciens types de chaque idiome lui donnait le moyen de remonter à des formes en quelque sorte primitives, susceptibles d'être analysées en leurs éléments composants; pour lui, un mot tel que sanskrit *émi*, grec *eîmi*, lituanien *eimi*, s'analysait naturellement en un mot *ai*, signifiant « aller » et un mot *mi*, signifiant « moi ». Mais, pour une analyse comme celle-ci, qui est plausible — quoique naturellement indémontrable — Bopp était amené à en proposer cent autres qui étaient ou peu vraisemblables ou tout à fait inadmissibles. Après une cinquantaine d'années d'essais infructueux de ce genre, on a compris que l'origine première des formes grammaticales est hors de nos

prises. Toutes les langues connues n'apparaissent qu'à une date plus ou moins basse, la plupart seulement à l'époque moderne et toutes sous des formes achevées, qui supposent un long développement antérieur ; les langues des peuples de civilisation inférieure ont des structures souvent délicates ; elles n'ont pas derrière elles une histoire moins longue que celles des langues des peuples les plus civilisés. Aucun idiome, quel qu'il soit, ne donne ni de près ni de loin l'idée de ce qu'a pu être une langue « primitive », et par suite, aucune donnée positive ne permet, non pas de résoudre, mais même d'aborder le problème de la première origine des formes grammaticales. Les linguistes étudient les transformations des systèmes grammaticaux ; ils ne s'occupent pas de la création de ces systèmes. Sans doute, quand un mot passe au rôle de forme grammaticale, on peut dire, en un certain sens, qu'il a été *créé* une forme, mais cette création a lieu à l'intérieur d'une langue qui offre déjà une organisation grammaticale complète, et cette création ne donne pas une idée de ce qui a pu se passer en un temps où il n'existait aucun commencement d'organisation grammaticale. Toutefois on peut retenir que, l'analogie étant par définition exclue de la première origine des formes, le seul procédé qui reste est l'attribution progressive d'un rôle grammatical à des mots autonomes ou à des manières de grouper les mots. En ce sens, Bopp avait évidemment raison ; mais son illusion était de croire que, avec les données tardives que l'on possède, on puisse même entrevoir la façon dont les formes existant dans les langues connues soit par des données historiques, soit par des procédés comparatifs, ont acquis pour la première fois les valeurs qu'elles possèdent.

Par le fait qu'ils renonçaient à déterminer l'origine première des formes et qu'ils se proposaient seulement d'en suivre le développement, les linguistes ont été conduits à s'attacher surtout à l'étude des innovations analogiques : car, étant donné un système une fois constitué qui se transforme peu à peu, l'analogie est le grand agent qui modifie sans cesse les détails, et même parfois la structure générale du système. Le mouvement linguistique qui a commencé vers 1870 et qu'on a souvent qualifié de

mouvement « néo-grammairien » est tout entier dominé par deux idées : la constance de ces correspondances entre phonèmes d'une même langue à deux dates successives, qui sont connues sous le nom de « lois phonétiques », et l'importance attribuée aux innovations analogiques. Quand Brugmann et Osthoff ont publié à partir de 1878 une collection de travaux conçus suivant les idées nouvelles, ils l'ont appelée *Morphologische Untersuchungen*, et l'analogie tenait la plus grande place dans le recueil. Les *Prinzipien der Sprachgeschichte* de M. H. Paul, qui ont été l'exposé des principes de l'école nouvelle, sont essentiellement une théorie de l'analogie. Et quand le regretté V. Henry a voulu répandre en France les idées des « néo-grammairiens », c'est par une étude sur l'*Analogie* qu'il a débuté.

Sans avoir jamais été perdu de vue, l'autre procédé d'innovation, le passage de mots autonomes au rôle d'agents grammaticaux, a été beaucoup moins étudié durant les quarante dernières années. On commence maintenant à s'y attacher de nouveau. L'importance en est en effet décisive. Tandis que l'analogie peut renouveler le détail des formes, mais laisse le plus souvent intact le plan d'ensemble du système existant, la « grammaticalisation » de certains mots crée des formes neuves, introduit des catégories qui n'avaient pas d'expression linguistique, transforme l'ensemble du système. Ce type d'innovations résulte d'ailleurs, comme les innovations analogiques, de l'usage qui est fait de la langue, il en est une conséquence immédiate et naturelle.

*
* *

Une phrase consiste essentiellement en une affirmation ; elle comprend au besoin un sujet, c'est-à-dire un mot qui indique de qui ou de quoi quelque chose est affirmé, et, normalement, un prédicat, c'est-à-dire un mot qui indique ce qui est affirmé. Le sujet peut n'être pas exprimé, s'il est connu des interlocuteurs ; c'est ce qui arrive à l'impératif, forme qui, par définition, s'adresse à un interlocuteur : *viens*, *venez* n'ont pas besoin de sujet ; ce n'est que si l'interlocuteur n'est pas assez désigné par la situa-

tion qu'on prévient la personne interpellée : *viens, Pierre*. Si l'on attend quelqu'un, on peut en russe dire : *pridèt* « il va venir » ou, en latin *uenit* « il vient » sans avoir besoin de désigner autrement la personne, qui est connue par la situation ou par une phrase antérieure. Ou bien, surtout dans les langues indo-européennes, la forme du verbe peut suffire à désigner la personne : russe *pridu* « je vais arriver », latin *uenio* « je viens ». Sauf les cas de ce genre, une phrase se compose essentiellement d'un sujet et d'un prédicat, le prédicat pouvant être un nom comme dans le russe *dom nov* « la maison est neuve », ou un verbe comme dans le russe *Pëtr pridët* « Pierre va arriver ». Ces mots sont les mots principaux de la phrase. Une phrase peut comprendre d'autres mots principaux, ainsi un complément dans une phrase comme *aedifico domum* « je bâtis une maison » ou *eo Romam* « je vais à Rome », ou *uenio Roma* « je viens de Rome », ou *habito Romae* « j'habite à Rome », en latin. Les mots principaux sont ceux qui indiquent les idées essentielles pour lesquelles est faite la phrase.

Mais on ne fait pas une phrase seulement avec des mots principaux. Il faut aussi le plus souvent d'autres mots qui déterminent, qui précisent la valeur de ces mots principaux. Soit une phrase comme : *laissez venir à moi les petits enfants*. L'orthographe française y distingue sept mots différents. Sans insister sur *à moi* où *à* n'est guère plus qu'un élément grammatical, ni provisoirement sur l'article *les* qui est aussi une sorte d'outil grammatical, on a ici deux mots dont chacun est groupé avec un autre : *laissez* et *petits*. Chacun de ces mots a un sens propre ; mais ce sens ne prend une valeur dans la phrase que par le groupement avec le mot voisin. *Laissez* peut être un mot principal, dans *laissez cela* par exemple ; mais ici *laissez venir* forme un ensemble, où *laissez* est, en quelque mesure, un auxiliaire. *Petit* a sa valeur propre, et est, même, en tant qu'un adjectif peut l'être, le mot principal, par exemple si l'on dit : *apportez le petit paquet* (et non pas le gros paquet qui est à côté) ; mais ici *petits* n'indique qu'une qualité accessoire d'*enfants*. Outre les mots principaux, il y a donc des mots *accessoires*. Tel mot qui est principal

dans une phrase est accessoire dans une autre ; *venir* est accessoire dans la phrase exclamative : *il vient me dire cela !* Et il y a tous les degrés intermédiaires entre les mots principaux et les mots accessoires ; *laissez* dans la phrase citée est moins un mot accessoire que *faire* dans *faites le venir*. Mais, dans toute phrase donnée, il importe de bien marquer la distinction entre les mots principaux et les mots qui sont plus ou moins accessoires.

Or, de ce qu'un mot est accessoire, il résulte deux sortes d'altérations, les unes touchant le sens, les autres touchant la prononciation.

A chaque fois qu'un élément linguistique est employé, sa valeur expressive diminue et la répétition en devient plus aisée. Un mot n'est ni entendu ni émis deux fois exactement avec la même intensité de valeur. C'est l'effet ordinaire de l'habitude. Un mot nouveau frappe vivement la première fois qu'on l'entend ; dès qu'il a été répété, il perd de sa force, et bientôt il ne vaut pas plus qu'un élément courant depuis longtemps. Ceci est plus vrai encore d'un groupe de mots : la plupart des gens parlent et surtout écrivent au moyen de formules toutes faites, de « clichés » ; aussi deux mots usuels apparaissent-ils presque neufs s'ils sont rapprochés pour la première fois ou si, du moins, on les rapproche alors qu'ils ne sont pas rapprochés d'ordinaire ; Horace a signalé depuis longtemps ce que vaut une alliance de mots nouvelle, une *iunctura nova* ; il l'a montré par beaucoup d'exemples. Les orateurs et les écrivains qui se soucient de style, c'est-à-dire d'expression, s'efforcent avant tout de combiner les mots d'une manière qui ne soit pas banale et qui par suite puisse faire impression sur l'auditeur ou le lecteur. Et si, au bout de quelques dizaines d'années d'usage littéraire, une langue est en général usée par la littérature, si tous les écrivains dans toutes les grandes langues de l'Europe en sont maintenant presque réduits à écrire d'une manière ou banale ou artificielle, c'est sans doute en grande partie parce que le nombre des alliances de mots nouvelles qui sont pratiquement possibles dans un idiome donné est limité.

Si donc un groupement de mots devient fréquent, s'il est sou-

vent répété, il cesse d'être expressif, et il est reproduit de plus en plus automatiquement par les sujets parlants. Il y a eu un temps où *je laisse venir* a constitué deux mots vraiment distincts et où *laisser* a eu dans une expression de ce type toute sa valeur sémantique. Mais on a pris l'habitude de grouper *laisser* avec un infinitif, et il en est résulté un affaiblissement rapide de ce mot qui a perdu son sens propre, pour devenir une sorte d'auxiliaire d'un mot principal. Néanmoins *laisser* n'a pas encore passé au rôle d'élément grammatical parce que l'on exprime par là une notion très spéciale et encore concrète et que *laisser* garde une autonomie nette de sens et de forme.

De ce qu'un mot est groupé avec un autre d'une manière qui tend à devenir fixe dans certains cas, il résulte pour ce mot la perte d'une partie de son sens concret dans ces constructions. Soit par exemple le mot *pied* ; employé isolément, il désigne une partie du corps humain très définie, de forme très spéciale ; groupé avec le nom d'un objet, dans des expressions comme *le pied d'une table, d'une chaise, d'une lampe* ou *le pied d'une montagne*, le mot perd sa valeur concrète tout entière, et il n'en reste plus qu'un élément abstrait : partie d'un objet qui supporte et est en contact avec une surface portante. Ainsi que l'a montré M. Wundt, on ne doit pas parler de métaphore dans les cas de ce genre, comme on continue malheureusement de le faire ; le terme est impropre. Il s'agit vraiment d'un autre mot ; et, en russe par exemple, on emploie en ce cas non le mot *noga* « pied », mais le dérivé *nožka* pour désigner le « pied » d'un meuble. Inversement, le groupe de mots a souvent un sens plus précis, plus concret que ne le fait même attendre le rapprochement des mots composants. Quand on parle d'un *pied de lampe*, on ne pense presque plus ni à *pied*, au sens abstrait qui vient d'être défini, ni à *lampe*, mais à un objet d'aspect particulier qui porte ce nom ; on peut à la réflexion se représenter les deux éléments du groupe, on en a peut-être un sentiment vague ; mais, en somme, *pied de lampe* est l'équivalent d'un mot un désignant un objet un. Le groupement peut être tel que l'un des mots ne reçoive pas de caractéristique grammaticale propre : si en turc osmanli, on veut par-

ler d'un jardin du pacha qui n'est pas familier aux interlocuteurs, on dira *paša-nyn baγče-si* « le jardin du pacha », avec le signe *-yn* du génitif ; si l'on est dans une localité où le « jardin du pacha » est chose connue, le mot *paša* ne sera pas décliné, et l'on aura, en un groupe comportant une seule flexion, *paša baγče-si* ; il n'y a presque plus ici deux mots distincts, mais une locution d'ensemble. Dans les langues d'Extrême-Orient, telles que le chinois ou l'annamite, qui ne connaissent pas les affixes et où par suite on ne peut former de noms par dérivation, c'est en groupant deux mots dont chacun perd sa signification propre qu'on obtient des noms de personnes ou de choses ; soit par exemple en annamite *phép* « règle » et *toán* « compter » ; on aura un mot *phép toán* signifiant « calcul, arithmétique » ; avec *ban* « tirer », on aura *phép ban* « tir », etc. ; de même *thay* « maître » forme les noms de professions libérales ; si *phép thuoc* est la « médecine », *thay thuoc* est le « médecin » etc. En somme le groupement habituel ôte aux mots et leur force expressive et la force expressive de leur union et même leur valeur concrète propre.

Les mots ainsi groupés par le sens sont la plupart du temps juxtaposés les uns aux autres. Ils se comportent dès lors dans la prononciation à peu près comme un mot long. On sait que la façon dont les mots sont unis dans la prononciation courante des phrases ne répond pas à la coupe des mots telle qu'elle apparaît dans la graphie ordinaire des langues modernes. Il s'est établi la convention que tout élément séparable ayant un rôle propre dans la phrase est, dans l'écriture, isolé de tout autre élément ; cet usage clair et commode est fondé uniquement sur le rôle que les mots jouent dans la phrase et sur la manière dont ils s'y comportent, et ne tient aucun compte de la prononciation. Par exemple l'article français, qui ne saurait en aucun cas être employé seul et qui fait toujours partie d'un groupe de noms, est écrit isolément parce qu'il peut se séparer du substantif qu'il détermine, et qu'on peut dire : *les enfants*, *les petits enfants*, *les pauvres petits enfants* etc. Au point de vue de la prononciation, *les enfants*, *les petits*, *les pauvres* forment dans ces groupements chacun un seul mot. La définition du *mot phonétique* ne recouvre pas celle du *mot syntaxique*.

Le groupe de mots tend à constituer plus ou moins un mot phonétique unique.

Or, des observations nombreuses ont montré que les mêmes éléments sont prononcés d'une manière d'autant plus brève qu'ils font partie d'un mot plus long : en français l'*â* de *pâté* est beaucoup plus bref que celui de *pâte*, et l'*â* de *pâtissier* et surtout de *pâtisserie* est plus bref que celui de *pâté*. L'abrègement a des conséquences graves : le timbre des voyelles en est souvent altéré ; une voyelle abrégée tend à se fermer ou, si elle est déjà fermée, peut s'amuir tout à fait. Les mots accessoires groupés avec d'autres tendent de ce chef à s'abréger et à changer de prononciation. De plus, et par le fait de l'abrègement, et par le fait que, étant accessoires, ils sont prononcés sans effort et entendus sans attention spéciale, ils sont négligés, dénués d'intensité, ils ne sont plus articulés qu'à demi. L'histoire des langues montre que, par suite, les mots accessoires ont des traitements phonétiques aberrants. On a souvent invoqué contre le principe de la constance des « lois phonétiques » les traitements spéciaux que présentent les mots accessoires. L'argument ne porte pas, on le voit ; les mots accessoires se trouvent dans des conditions particulières qui déterminent des prononciations particulières : leurs éléments constituants, étant abrégés et faiblement articulés, sont exposés à s'affaiblir ou à disparaître dans des cas où les éléments d'un mot principal subsistent intacts ou subissent des modifications tout autres. On a autrefois utilisé par exemple contre la règle de la constance des « lois phonétiques » le fait que le *th* anglais est demeuré sourd (*th* dit dur) à l'initiale de presque tous les mots de la langue, mais qu'il est devenu sonore (*th* doux) dans l'article *the* ; on sait maintenant que les consonnes sourdes initiales des mots accessoires sont sujettes à subir des affaiblissements propres à ces mots ; outre l'anglais, le phénomène s'observe par exemple en irlandais, en scandinave et en arménien ; et il n'est pas particulier aux langues indo-européennes ; on l'a retrouvé en polynésien, à Samoa par exemple.

Les altérations phonétiques subies par les mots accessoires sont parfois très profondes. Si l'on n'avait pas la forme du gotique

himma daga « ce jour-ci, aujourd'hui », on pourrait avoir peine à croire qu'un ancien *hiu tagu* « ce jour-ci » soit devenu en vieux haut allemand *hiutu* (allemand moderne *heute*) « aujourd'hui », et que *hiu dagu* « ce jour-ci » soit devenu en vieux saxon *hiudu* « aujourd'hui » ; les doutes qui pourraient subsister sont levés par le passage parallèle de *hiu jâru* « cette année » à *hiuru* (allemand moderne *heuer*) ou par celui de *hînaht* « cette nuit » à *hînet* en moyen haut allemand, à *heint* « aujourd'hui » en bavarois moderne. L'accent portant sur le début du composé, sur le démonstratif qui renferme l'essentiel de l'idée, à savoir l'indication qu'il s'agit de « ce qui est le plus près », tout le reste du mot a été réduit à presque rien et est devenu méconnaissable. — Les mots accessoires en viennent à ne plus ressembler aux mots principaux là même où, à l'origine, ils étaient identiques. Ainsi dans un ancien dialecte grec, le béotien, les mots en *-a-* long ont un génitif pluriel en *-âôn*, sans contraction ; mais l'article correspondant contracte *-âôn* en *-ôn*, et l'on a *tôn drakhmâôn* par exemple : au lieu d'une flexion une, il y a deux flexions distinctes, l'une pour les mots principaux, l'autre pour l'article.

L'affaiblissement du sens et l'affaiblissement de la forme des mots accessoires vont de pair ; quand l'un et l'autre sont assez avancés, le mot accessoire peut finir par ne plus être qu'un élément privé de sens propre, joint à un mot principal pour en marquer le rôle grammatical. Le changement d'un mot en élément grammatical est accompli.

*
* *

La constitution de formes grammaticales par dégradation progressive de mots jadis autonomes est rendue possible par les procédés qu'on vient de décrire sommairement, et qui consistent, on le voit, en un affaiblissement de la prononciation, de la signification concrète des mots et de la valeur expressive des mots et des groupes de mots. Mais ce qui en provoque le début, c'est le besoin de parler avec force, le désir d'être expressif. L'histoire de la négation fournit une illustration de ce principe.

La négation s'exprimait en indo-européen commun par un petit mot accessoire *ne* qui subsiste clairement dans le *na* du sanskrit, le *ne* du slave et du lituanien, le *ni* du gotique par exemple. Mais ce petit mot très bref, qui tendait à être inaccentué et qui en lituanien et en russe en est venu à se grouper étroitement avec le mot principal sur lequel il porte, est devenu rapidement très inexpressif. Là où l'on avait besoin d'insister sur la négation — et les sujets parlants éprouvent presque toujours le besoin d'insister, car on parle le plus souvent pour agir sur les autres en quelque manière, et l'on fait ce qu'il faut pour les toucher —, on a été conduit à renforcer la négation *ne* par quelque autre mot. C'est ce qui est arrivé en latin par exemple ; de même que pour dire « non », l'allemand en est venu à dire « pas un » à savoir *nein*, le latin ancien a dit *noenum* « pas un », au lieu de *ne*. Étant un mot accessoire, *noenum* a subi un traitement particulier et a abouti à *nôn*. Mais dès lors on ne retrouvait plus « pas un » dans *nôn*, et, à l'époque historique, le *nôn* latin n'est pas sensiblement plus expressif que le *na* sanskrit, le *ne* slave, le *ni* gotique. Le français a été, par suite, conduit à renforcer sa négation *ne*, issue de *nôn* par un traitement particulier de mot accessoire, au moyen de petits mots, tels que *pas*, *point*, *mie*. On sait comment *pas* a perdu, dans les phrases où il était un accessoire de la négation, tout son sens propre — sens conservé parfaitement dans le mot isolé *pas* —, comment dès lors, *pas* est devenu à lui seul un mot négatif, servant à exprimer la négation, et comment, par suite, le *pas* français n'est plus expressif à son tour et appelle un nouveau renforcement par des mots accessoires ; on est amené à dire *pas du tout*, *absolument pas*, ou à recourir à des tours tout nouveaux, comme l'exclamation argotique actuelle : *tu penses*, *s'il est venu !* manière fortement expressive de dire : « il n'est pas venu ». — L'histoire de la négation allemande *nicht*, qui étymologiquement signifie « pas une chose », est parallèle à celle des négations latine et française. — Les langues suivent ainsi une sorte de développement en spirale : elles ajoutent des mots accessoires pour obtenir une expression intense ; ces mots s'affaiblissent, se dégradent et tombent au niveau de simples

outils grammaticaux; on ajoute de nouveaux mots ou des mots différents en vue de l'expression; l'affaiblissement recommence, et ainsi sans fin.

Les catégories grammaticales qui sont sujettes à être exprimées au moyen de mots devenus éléments grammaticaux sont, par suite de ce qui vient d'être dit, celles qui ont un certain caractère expressif; c'est du moins ce qui a lieu dans les langues où il existe des formes grammaticales caractérisées par des affixes soudés aux mots. Les langues, qui, comme le chinois et plus encore peut-être l'annamite, ignorent toute affixation, sont conduites à exprimer plus de catégories par des mots accessoires qu'on appelle souvent mots vides. Mais, dans les langues indo-européennes, les catégories qui, comme celles du présent et de l'aoriste dans le verbe, expriment simplement des faits, sont peu propres à être caractérisées par des mots accessoires progressivement soudés. Ou si des mots accessoires y interviennent, ce n'est que secondairement et quand le procès d'affaiblissement est achevé, comme on le verra. Il en va autrement des catégories qui ont une signification plus intense et qui appellent par suite une expression nette et fortement caractérisée. L'histoire du parfait et celle du futur sont instructives à cet égard.

On entend par parfait une catégorie à sens très fort, qui indique l'action en tant qu'elle est achevée, le résultat acquis de l'action, et non pas l'action elle-même dans son développement et sa durée ou l'acte pur et simple. Un sens de ce genre était exprimé en indo-européen par un type tout particulier, dont les désinences et le vocalisme radical étaient propres à ce type, qui de plus présentait dans une grande partie des cas un redoublement d'une partie initiale de la racine, ainsi en grec *léloipa* « j'ai laissé », etc. Cette formation, très à part et par suite très expressive, ne s'est pas maintenue au cours du développement historique des langues indo-européennes, en partie parce que le sens s'est affaibli et dégradé jusqu'à celui du présent, comme dans *memini* « je me souviens », du latin ou *man* « je pense » du gotique, ou

jusqu'à celui du prétérit, comme dans *cecini* « j'ai chanté » du latin ou *haihald* « j'ai tenu » du gotique, en partie parce que la structure du type était trop spécifiquement indo-européenne pour survivre à la période ancienne de l'indo-européen : les changements généraux qui se sont produits dans le système ne laissaient pas subsister les conditions de formation qu'exige le parfait indo-européen.

Mais la forme, en disparaissant, laissait un vide. Car on éprouve le besoin de bien marquer l'action achevée dont on envisage le résultat. On y parvient le plus souvent par le groupement d'une forme nominale rattachée à un verbe accessoire. Et ceci se conçoit bien : l'action achevée n'est déjà plus un procès, mais une chose ; elle est donc propre à être exprimée par un nom plutôt que par un verbe, puisque le propre du verbe est d'exprimer un procès. Cette tendance est surtout manifeste au passif, et c'est pour cela que les formes composées qui servent à l'expression du parfait se présentent d'abord le plus souvent sous la forme passive : le latin a *dictus est* « il (quelqu'un) a été dit », et surtout sous forme impersonnelle *dictum est* « il (quelque chose) a été dit, on a dit », dès une époque antérieure à l'époque historique (le procédé est du reste commun non seulement au latin et à l'osco-ombrien, mais aussi au celtique, ce qui atteste une antiquité relativement grande de l'innovation).

Ce n'est que beaucoup plus tard, à l'époque où se sont constituées les langues romanes, qu'un type actif, de structure toute différente, apparaît : *habeo dictum* « j'ai dit ». Quand ce type roman s'est constitué, il avait une grande force expressive : je possède quelque chose qui est dit. Le procédé, très frappant, se retrouve en germanique, après la période la plus ancienne de la langue (il n'y en a pas encore trace en gotique au IV^e^ siècle après Jésus-Christ), sans doute par une imitation d'une manière de dire latine qui semblait frappante et commode ; de ce qu'il y a ici une imitation d'une manière de grouper les mots, on ne conclura pas que le germanique a emprunté au latin une forme grammaticale : les formes grammaticales proprement dites ne semblent guère s'emprunter ; et, au moment où l'imitation a pu

avoir lieu, le type *habeo dictum* comportait sans doute encore deux mots sentis comme nettement distincts : ce n'était pas encore une forme grammaticale, mais un groupement de mots. Avec le temps, le type *j'ai dit* s'est unifié, et, de bonne heure, en français, c'est purement et simplement une manière d'exprimer l'action accomplie, où l'on ne reconnaît plus la valeur ni de *ai* ni de *dit* ; mais alors le groupe cesse aussi d'être expressif ; il perd sa valeur de parfait pour devenir un simple prétérit ; et, comme il fait alors concurrence au prétérit simple, *je dis*, qui est beaucoup moins clair, de formation plus compliquée et en partie spéciale, fléchi d'ailleurs d'une manière particulière (*nous dîmes, vous dites, ils dirent*), parfois enfin ambigu (*je dis, il dit, vous dites* servent à la fois pour le présent et pour le prétérit), le prétérit simple a tendu à disparaître : il est aujourd'hui entièrement sorti de l'usage à Paris et dans toute la région où le français se parle à la manière parisienne, c'est-à-dire dans un rayon de deux à trois cents kilomètres autour de Paris. L'affaiblissement progressif de la valeur du type *j'ai dit* a abouti à en faire un simple prétérit, sans aucun reste de la valeur de parfait. Le cycle est désormais parcouru, et, pour se donner un parfait, le français devra recourir à quelque tour nouveau, dont on n'entrevoit pas encore la naissance.

Des phénomènes analogues à ceux que l'on observe dans l'histoire du latin et des langues romanes, et notamment du français, ont eu lieu dans beaucoup d'autres langues, et d'une manière indépendante. Par exemple, le parfait indo-européen était déjà presque sorti de l'usage en perse à l'époque de Darius, donc dès la fin du VI^e^ siècle av. J.-C. ; on y suppléait par une forme nominale de type passif, qui avait une valeur très nette de parfait ; pour indiquer ce qui est accompli, on disait *ima tya maná kṛtam* « voici ce qui a été fait par moi », *kṛtam* signifiant ici « fait ». Mais il subsiste encore en vieux perse des prétérits simples pour exprimer l'acte, et l'aoriste *akumâ* signifie « nous avons fait », là où il s'agit d'exprimer le fait pur et simple. Par la suite, le prétérit simple s'est éliminé ; il n'est resté que l'ancienne forme composée ; mais elle a pris le caractère actif, et *kard* signifie en persan

« il a fait » ; *man* dans *man kardam* a pris la valeur d'un cas sujet ; on fléchit *kardam* « j'ai fait », qui produit l'impression d'une forme une et où l'on ne reconnaît plus en rien l'ancien participe ; pour le sens, *kardam* est un simple prétérit et n'a plus la signification du parfait, pas plus que le français moderne *j'ai fait*.

Sous une forme tout autre, les langues slaves offrent un développement exactement parallèle. A date ancienne, on y trouve un prétérit simple tel que *budixŭ* « j'ai éveillé » et un parfait composé : *budilŭ jesmĭ* « j'ai éveillé » avec (valeur de parfait), littéralement « je suis l'éveilleur ». La valeur du parfait n'apparaît déjà plus qu'affaiblie à l'époque des plus anciens textes, et, dans les langues slaves modernes un type tel que *voz-budil* en russe, ou *wz-budzikem* en polonais est un type simple, à valeur pure et simple de prétérit. Dans plusieurs des principales langues slaves, notamment en russe et en polonais, il y a même plus d'autre prétérit que l'ancien parfait composé ; le prétérit simple est sorti de l'usage depuis longtemps.

L'évolution est donc la même dans plusieurs langues, et il ne serait pas difficile d'ajouter d'autres exemples à ceux qu'on vient de passer en revue. Le parfait tend à s'exprimer par des formes composées, de caractère nominal ; aussitôt entrées dans l'usage courant, ces formes tendent à perdre leur valeur de parfait ; elles se dégradent au niveau de simples prétérits ; du même coup elles perdent leur caractère nominal et apparaissent comme des formes verbales. Alors, étant en général plus régulières que les anciennes formes du prétérit simple, elles tendent à remplacer celles-ci.

Le futur ou le présent duratif donneraient lieu à des remarques du même ordre que celles qui viennent d'être présentées à propos du parfait

Même dans les langues qui comportent une conjugaison compliquée, il peut n'y avoir aucune expression des actions futures par des formes grammaticales particulières. Tel est le cas des langues sémitiques et d'une grande partie des anciennes langues indo-européennes par exemple. L'ancien germanique n'avait pas

de futur, et aujourd'hui encore on peut à peine dire que l'allemand ait un futur. Pour indiquer l'action à venir, on recourt souvent à des formes qui indiquent qu'on a l'intention de faire quelque chose; la forme grammaticale qui, dans l'indo-européen commun d'où sont issues toutes les langues indo-européennes, avait ce sens, était le mode subjonctif; et il se trouve ainsi que, en latin par exemple, des formes comme *erit* ou *dicet* qui, de par leur origine, sont des subjonctifs, ont pris la valeur de futur et n'ont même plus d'autre valeur en latin à l'époque historique; seule, la grammaire comparée avertit que *erit* et *dicet* du latin ont été, à une époque préhistorique, des subjonctifs. A une époque plus récente du développement de certaines langues indo-européennes, le verbe *vouloir* est devenu un auxiliaire servant à l'expression du futur; dans la bouche de beaucoup de Français, *je veux faire* forme déjà un groupe dont le sens s'est assez affaibli pour équivaloir à une sorte de futur. Dans l'anglais *I will make*, le développement est plus avancé encore. Dans les langues slaves du Sud et en grec moderne, l'expression du futur est obtenue au moyen d'un verbe signifiant « vouloir », mais si dégradé pour la forme et pour le sens que seul le linguiste peut maintenant l'y reconnaître. En grec moderne par exemple *thelô ina* « je veux que », réduit à *thelô na*, *thena* et simplement *tha*, ne laisse plus rien deviner du verbe signifiant « vouloir ». Ailleurs on peut se servir de mots signifiant « devoir », comme dans l'anglais *I shall make* ou dans l'arménien moderne occidental *bidi anem* « je ferai », littéralement « il y a nécessité que je fasse »; et alors le mot accessoire peut se réduire et devenir méconnaissable; au lieu de *bidi*, on trouve simplement *di* dans nombre de parlers arméniens. C'est par ce moyen que les langues romanes se sont donné un futur quand le futur du latin ancien est devenu trop faible, trop inexpressif : *facere habeo*, qui est l'original de (*je*) *ferai*, signifie « j'ai à faire », c'est-à-dire « je dois faire ». L'infinitif et le verbe signifiant « avoir » se sont soudés; *avoir* a pris une flexion propre différente de celle du verbe isolé; en disant *je finirai* un Français ne pense pas à *finir* ni à *ai*; en disant *nous finirons*, il ne peut pas penser à *finir* ni à *-ons*, et il ne saurait être

question d'analyser *j'aimerai* ou *je viendrai*. Par suite, ces formes ont perdu toute valeur expressive. Comme on est le plus souvent tenté de parler de l'avenir avec quelque expression particulière, celle du désir, de l'attente, de la nécessité, le français en est venu à se refaire de nouveaux futurs qui aient encore une force d'expression : un futur prochain : *je vais faire*, où *je vais* n'est déjà plus qu'un auxiliaire et où le sens d'*aller* n'est plus perceptible ; *je veux faire*, où le sens de *vouloir* est encore bien perceptible ; *je dois faire*, où *je dois* n'est guère plus qu'un auxiliaire, mais où le sens de nécessité est net ; *j'ai à faire* ; etc. Le futur n'est pas une forme nécessaire ; mais dans les langues où il existe, il se refait constamment.

Là où il y a une conjugaison qui exprime le temps, il faut une forme propre à exprimer une action qui a lieu présentement. Mais si l'on veut insister sur la durée de l'action, on se sert souvent d'expressions complexes qui en viennent ensuite à fournir des formes simples. Le type « je suis faisant », *I am making*, de l'anglais est un exemple actuel de ce fait. Dans tous les parlers actuels de l'arménien, le présent s'exprime par des formes qui n'ont plus aucune valeur expressive particulière, mais qui ont été des formes composées exprimant la durée. Le type *sirum em* de la plupart des parlers de l'Arménie russe a signifié « je suis à aimer » et le type *g sirem* de la plupart des parlers de l'Arménie turque est le résultat de l'altération d'une forme complexe : *kay ew sirê* « il se tient et il aime » où *kay ew* s'est réduit à *ku* et *ku* à *kə* (*ə* désignant une sorte d'*e* muet) ; la prononciation *g* de l'ancien *k* est une particularité du parler arménien de cette région. L'expression *kay ew sirê* a exprimé très fortement la durée de l'action ; la forme moderne *g sire* ne vaut pas plus que le français *il aime* et n'est pas sensiblement moins une au sentiment des sujets parlants.

Les exemples de cette sorte pourraient être multipliés · toujours le besoin d'expression fait créer des groupes qui, par l'usage, perdent leur valeur expressive et servent alors de formes grammaticales, dénuées de force.

Dès lors on voit combien peu il est légitime de parler de

langues *synthétiques* et de langues *analytiques*. Ce n'est pas pour analyser qu'on emploie des formes composées ; c'est en vue de l'expression ; et ce n'est pas pour synthétiser qu'on a des formes unes : les formes unes résultent du rapprochement qui a lieu en fait entre mots groupés d'une manière habituelle. Quand on veut s'exprimer avec force, on donne à chaque notion une expression séparée ; on ne dit pas « je ferai », mais « j'ai la volonté de faire » ou « il faut que je fasse » ou « je suis sur le point de faire » ; il ne s'agit pas ici de logique, mais de sentiment à rendre et d'action à exercer sur un interlocuteur. Et si *je veux faire, je dois faire, je vais faire* n'expriment plus nettement la volonté, la nécessité, la proximité, c'est que du fait du groupement ordinaire, les mots *veux, dois, vais* ont perdu leur sens propre, leur valeur expressive, et qu'ils sont devenus de simples auxiliaires en attendant qu'ils fassent corps avec l'infinitif suivant. Le Romain qui disait *facere habeo* ne faisait pas d'analyse, pas plus que le Français qui dit *je ferai* ne fait de synthèse. Analyse et synthèse sont des termes logiques qui trompent entièrement sur les procès réels. La « synthèse » est une conséquence nécessaire et naturelle de l'usage qui est fait de groupes de mots.

Les mots ne sont du reste pas seuls à être sujets à devenir des éléments grammaticaux ; la façon de grouper les mots peut aussi devenir un procédé d'expression grammaticale. En latin où le rôle grammatical de chaque nom est indiqué par la forme de ce nom, il n'y a pas d'ordre nécessaire ; on peut dire pour « Pierre bat Paul » : *Petrus Paulum caedit*, ou *Paulum Petrus caedit*, ou *caedit Petrus Paulum*, ou *caedit Paulum Petrus*, etc. L'ordre n'est pas indifférent ; il sert à indiquer certaines nuances : suivant qu'on met en avant *Petrus* ou *Paulum*, on attire l'attention sur l'un ou l'autre mot ; mais l'ordre n'indique en rien le rôle grammatical des mots. En français ou en anglais au contraire, c'est la place respective des mots qui indique leur rôle, et en interchangeant la place de *Pierre* et de *Paul* dans *Pierre bat Paul*, on changerait aussi les rôles grammaticaux des deux noms. Ici, un ordre de mots devenu habituel pour quelque raison a pris le caractère de « morphème », c'est-à-dire de marque d'une catégorie gram-

maticale. La valeur expressive de l'ordre des mots, que l'on observe en latin, a été remplacée par une valeur grammaticale [1]. Le phénomène est de même ordre que la « grammaticalisation » de tel ou tel mot; au lieu que ce soit un mot employé en groupe avec d'autres qui prenne le caractère de « morphème » par un effet de l'habitude, c'est une manière de grouper les mots. Ici encore, il y a vraiment création d'outils grammaticaux nouveaux, et non pas transformation. Tout le parti que le français et l'anglais tirent de l'ordre des mots pour marquer les relations des parties de la phrase entre elles est une création de ces langues : ni le latin ni l'ancien germanique n'offraient rien de pareil.

1. Sur le rôle du sentiment dans la création des formes grammaticales, on verra les livres de M. Bally : la *Stylistique*, le *Précis* et le *Traité* et *Le langage et la vie*, et, plus récemment le recueil de M. Leo Spitzer, *Aufsätze zur romanischen Syntax und Stylistik* (Halle, 1918).

SUR LA DISPARITION

DES

FORMES SIMPLES DU PRÉTÉRIT [1]

On observe en phonétique des tendances générales qui se manifestent au cours du développement des langues les plus diverses. M. Grammont a mis en évidence de pareilles tendances dans ses travaux sur la dissimilation et sur la métathèse. Et ce ne sont là que des exemples ; car on peut constater des « lois » générales dans toutes les parties de la phonétique. Il est beaucoup plus malaisé de saisir et de formuler en morphologie des tendances générales parce que les faits morphologiques ont un caractère singulier — au sens propre du mot — et en quelque sorte fortuit. Conditionnés par des données anatomiques et physiologiques invariables et qui sont sensiblement identiques dans toutes les sociétés, les faits phonétiques se meuvent dans un cadre étroit, et par suite on retrouve partout des situations comparables entraînant des évolutions pareilles ou du moins très analogues les unes aux autres. En morphologie, la liberté est beaucoup plus grande ; la grammaire de deux langues de famille distincte diffère du tout au tout ; et, à l'intérieur d'une même famille, les grammaires de deux langues deviennent souvent très différentes au bout de quelques siècles d'évolution divergente : on sait combien diffèrent déjà les grammaires des diverses langues romanes. Il résulte de là que les tendances générales n'apparaissent pas au premier abord, et que, là où l'on en aperçoit, elles se laissent très malaisément formuler. Cependant, au moins à l'intérieur d'une même famille où certaines situations analogues

1. Germanisch-Romanische Monatsschrift, I (1909), p. 521 et suiv.

se rencontrent, on peut apercevoir des tendances morphologiques générales. C'est un cas de ce genre qui va être indiqué.

Dans le français courant de Paris et dans un vaste rayon autour de Paris, le prétérit simple dit passé défini, du type *il eut, il fut, nous crûmes, il partit,* est entièrement sorti de l'usage parlé. La forme s'écrit encore, sinon dans le style courant où elle tend à disparaitre, du moins dans des textes didactiques dont la langue est traditionnelle et artificielle ; la grammaire enseigne à l'employer dans le récit, où la forme du passé composé serait contraire à l'usage classique ; mais cet emploi ne répond plus au sentiment actuel des sujets parlants. L'Atlas linguistique de MM. Gilliéron et Edmont montre que, dans tout le français proprement dit, le prétérit simple est une forme morte ; un coup d'œil jeté sur les cartes 97, 338, 360, 976, 1154 suffit à l'indiquer immédiatement. Au nord de Paris, pas trace de « passé défini » ; au Sud, il faut aller jusqu'aux départements de l'Allier, de la Creuse et de la Vienne pour commencer à en trouver quelques-uns sur les cartes ; à l'Ouest, la limite est plus proche : le prétérit simple existe encore en Normandie ; de là vient peut-être que Guy de Maupassant a beaucoup employé cette forme grammaticale. Mais, là même où le prétérit simple subsiste encore, la forme composée semble dominer déjà, ou du moins fait concurrence à la forme simple. L'Atlas linguistique signale des traces de prétérit simple sur plusieurs points de l'Allier ; mais à Moulins (Allier) où je suis né et où j'ai été en partie élevé, la forme n'est pas dans l'usage courant plus qu'à Paris ; mes grands-parents maternels, nés à Moulins en 1817, ne l'employaient pas ; ma grand'mère maternelle que je pouvais encore observer en 1909 (quand cet article a été écrit) n'y recourait jamais ; et je n'ai pas eu occasion de l'entendre chez les paysans des environs de Moulins, qui, il est vrai, ne parlent pas patois et qui sont sous l'influence du français urbain. Pour moi, qui ai été élevé dans les départements du Cher et de l'Allier, et qui me suis fixé ensuite à Paris, la forme du « passé défini » m'apparaît comme barbare ou pédante, et je ne puis l'entendre dans la conversation ou la lire dans une lettre familière sans en

être vivement choqué. A vrai dire on ne l'entend que chez des personnes originaires des parties de la France où le prétérit simple subsiste, surtout chez les méridionaux (car le passé simple subsiste dans les parlers du Midi de la France) ou chez des personnes qui ont subi trop fortement l'influence de la langue écrite. Tel littérateur qui se sert du passé défini, a de fausses formes, ainsi les frères Rosny quand ils écrivent, à la fin de leur *Daniel Valgraive*, « il se *dissolva* dans les ténèbres [1] ». La substitution de l'imparfait au passé défini, que M. Lanson, dans son *Art de la prose*, p. 266, donne comme un effet de style des romanciers naturalistes, provient sans doute en grande partie d'une répugnance qu'avaient ces écrivains à user d'une forme sortie de l'usage courant. Cette disparition doit être assez ancienne puisque, dès le début du XIX^e^ siècle, la forme ne figure plus dans le parler de personnes élevées dans une ville aussi méridionale que Moulins. Le français canadien, quoique reposant en partie sur le parler des Normands, ne la possède pas non plus (voir l'article de M. Meyer-Lübke dans cette revue I, p. 137). Parmi les autres langues romanes, on cite le rhétoroman comme étant en train de perdre le prétérit simple ; beaucoup de parlers n'en ont plus aucun reste (voir Meyer Lübke, *Grammaire des langues romanes*, II, 268). M. Mario Roques me signale aussi que, en roumain, le prétérit simple sort actuellement de l'usage parlé.

Le prétérit simple indicatif tend de même à s'éliminer en allemand ; on peut voir le détail des faits dans un récent article de M. Jacki, *Das starke Präteritum*, *Beiträge* de Paul et Braune, XXXIV, 425-597. La disparition est dès maintenant un fait accompli dans tout le Sud-Ouest du domaine allemand : Suisse, Alsace, Lorraine allemande jusqu'à la Moselle, Bade, Würtemberg ; il ne subsiste que des traces du prétérit dans le Palatinat bavarois et la Hesse rhénane, à savoir *war* ou quelque fois *hatte*, *ward*, *wollte* ; jusque dans le Luxembourg, le Nassau, le Sud de la Hesse supérieure, le prétérit simple n'est pas d'emploi courant.

1. J'ai lu récemment dans un journal, qui n'est en général pas mal écrit : *il extraya* (juillet 1920).

Plus à l'ouest, sauf *war*, le prétérit manque en Bavière et en Autriche. Un emploi libre et normal ne se rencontre plus que dans le Nord-Est ; Thuringe, Saxe, Silésie, etc. Même dans les régions où le prétérit simple existe encore, la différence de sens entre le prétérit simple et la forme composée n'est plus perçue en général, d'après les faits cités par M. Jacki. On notera à ce propos la curieuse remarque suivante de M. et Mme Stern, dans leur *Kindersprache*, p. 223 : « das Imperfekt ist entsprechend seiner viel grösseren Seltenheit in der Umgangssprache erst eine sehr späte Eroberung des Kindes und bleibt auch weiterhin abgesehen von einigen Hilfszeitwörtern, *war*, *hatte*, *wollte*, eine sehr spärlich gebrauchte Form » ; or, les observations de M. et Mme Stern ont été faites à Breslau, dans une région où le prétérit simple n'est pas sorti de l'usage.

On a essayé d'expliquer cette élimination du prétérit simple de l'allemand par des faits phonétiques propres aux parlers où le phénomène s'est produit ; mais M. Wunderlich a réfuté cette hypothèse, et M. Jacki a confirmé la valeur de sa critique (l. c. p. 254) [1]. C'est bien en tant que forme du passé que le vieux prétérit germanique tend à sortir de l'usage et à être remplacé par une forme composée. Cette tendance à la disparition du prétérit simple se marque sur certains domaines allemands dès le commencement de la période moderne (v. Wunderlich, *Der deutsche Satzbau*, 2° Aufl., I. 214 et suiv.).

Le slave présente des faits exactement semblables. Le slave

1. Répondant à cette observation, dans la *Germanisch-Romanische Monatschrift*, II, p. 383 et suiv., M. H. Reis a insisté sur le fait que les domaines où *er spielte* s'est confondu phonétiquement avec *er spielt* sont aussi, en gros, ceux où le prétérit du type *er spielte* a disparu. Il est donc probable que la confusion des formes a réellement aidé à la ruine du prétérit simple. En français aussi, on a invoqué l'ambiguïté des formes telles que *je finis* ou *je dis* pour expliquer la disparition du prétérit simple. Mais, en ce qui concerne le français, cette ambiguïté n'existe que dans une petite partie des formes : les types *j'aimai*, *je rendis*, *je reçus*, *je fus* ne sont pas ambigus ; et, même pour le germanique, les formes des verbes forts et toutes celles du pluriel n'étaient pas ambiguës. Du reste les faits observés dans d'autres langues montrent que ce facteur n'est ni le seul, ni sans doute le principal.

commun avait deux types de prétérits, l'un simple dit aoriste, type *nĕsŭ* « j'ai porté », *vŭzbudixŭ* « j'ai éveillé » etc. ; l'autre composé, type *neslŭ jesmĭ* « j'ai porté », *vŭzbudilŭ jesmĭ* « j'ai éveillé ». Il y a trace des deux dans tous les principaux dialectes. Et en vieux slave l'aoriste est beaucoup plus employé que la forme composée. Mais de très bonne heure, on voit l'aoriste s'éliminer. En vieux russe l'aoriste disparaît dès le moyen âge, si bien qu'il n'en subsiste aucune forme ayant valeur de prétérit. De même en polonais. D'autres langues ont conservé l'aoriste plus longtemps. Mais en serbe par exemple, l'aoriste qui s'était longtemps maintenu et qui figure encore dans la langue littéraire, sort actuellement de l'usage dans nombre de parlers populaires ; dans deux des grands groupes du serbe, le groupe de *ča* et celui de *kaj*, l'aoriste est déjà perdu ; et même dans le groupe de *što*, il tend presque partout à disparaître aussi (v. Rešetar, *Der štokavische Dialekt*, col. 192). En slovène, l'aoriste ne se rencontre que dans les monuments de Freising, puis à l'état de traces dans les plus anciens textes proprement slovènes connus à partir du XV^e^ siècle ; il a disparu entièrement par la suite (v. Vondrak, *Vgl. slav. Gramm.*, II, p. 154)[1].

Les formes telles que l'aoriste et le parfait, si fréquentes en iranien ancien, ne laissent en pehlvi pas même une trace ; le passé est exprimé en moyen iranien au moyen d'une forme qui repose sur l'adjectif verbal en *-ta* ; cette forme a commencé à se constituer de très bonne heure, comme on le voit par les inscriptions perses du VI^e^ siècle av. J.-C. L'iranien moderne a entièrement perdu l'ancien aoriste, et il faut aller jusque dans le Pamir pour trouver encore en usage l'ancien imparfait qui persiste dans un dialecte (v. Kuhn u. Geiger, *Grundriss der iranischen philologie*, I, 2, p. 340).

Quant à l'Inde, la substitution d'une forme nominale aux anciennes formes simples au cours de l'histoire des langues aryennes

1. Le fait que, à l'aoriste slave, la 2^e^ et la 3^e^ personne du singulier n'avaient qu'une seule forme causait une ambiguïté choquante et constituait une anomalie ; ceci a pu contribuer à faire éliminer cet aoriste.

se réfléchit déjà dans les textes sanskrits (v. J. Bloch, *Mémoires de la Société de linguistique*, XIV, 93 et suiv. [1]).

L'aoriste, qui s'est en général conservé en arménien, a été cependant éliminé dans les parlers arméniens des colonies d'Autriche-Hongrie (v. Adjarian, *Classification des dialectes arméniens*, p. 80). Et la forme composée du prétérit a pris une grande extension dans tous les parlers arméniens modernes.

Dans le procès de disparition dont on vient de voir des exemples, il y a deux moments à distinguer :

1° Création d'une forme composée de prétérit.

2° Généralisation de cette forme aux dépens du prétérit simple.

De ces deux moments, le second n'est constaté que dans une partie du domaine indo-européen ; le premier au contraire est général. Presque partout on voit se créer une forme composée du prétérit, qui existe d'abord concurremment avec le prétérit simple, et qui exprime le résultat acquis par une action donnée, qui ainsi tient la place de l'ancien parfait. Sauf les anciennes langues indo-iraniennes et le grec ancien, le parfait indo-européen ne s'est en effet maintenu nulle part à l'état de forme distincte, pas même dans les textes les plus anciens de chaque langue. C'est une forme composée différente dans chaque domaine, qui en tient lieu. Le grec moderne a par exemple ἔχω δεμένο (*exo demeno*) « j'ai lié », avec le verbe « avoir » et l'ancien participe parfait ; l'aoriste ἔδεσα (*edesa*) est le temps historique. L'arménien ancien a *sireal em* « j'ai aimé », à côté de la forme historique, *sireçi*, et les divers parlers arméniens modernes ont ces mêmes formes ou des formes équivalentes. Le vieux slave a *neslŭ jesmĭ* « j'ai porté » concurremment avec la forme historique *nèsŭ*, et plusieurs langues slaves conservent encore cette dualité de temps. A date ancienne, les langues italiques, celtiques et germaniques n'ont pas encore des formes composées pour le parfait, ou du moins n'en ont qu'au passif ; ceci tient à ce que le prétérit simple résulte de

1. Voir maintenant aussi J. Bloch, *La formation de la langue marathe*, p. 249 et suiv.

la combinaison de l'aoriste et du parfait indo-européens, et par suite retient quelque chose du sens du parfait. Ce n'est qu'à l'époque romane que s'est développé le type *j'ai aimé*. Et, en germanique, le gotique ignore encore ce type. Il n'y a d'ailleurs eu là qu'un retard, au moins en ce qui concerne l'italique et le germanique. Les dialectes celtiques qui ont recours à d'autres procédés pour exprimer le parfait n'ont pas développé en général de formes composées au moyen d'un participe et d'un auxiliaire.

Les deux types de prétérits, le type simple et le type à participe et auxiliaire, peuvent se maintenir indéfiniment côte à côte, avec leur différence de sens initiale. Mais la forme simple qui est ancienne et qui, malgré de nombreuses innovations analogiques, ne parvient jamais à n'avoir qu'une caractéristique unique pour tous les verbes, qui surtout a des aspects très divers dans les verbes forts, et qui enfin a souvent une flexion très singulière (ainsi en français, *nous aimâmes, vous aimâtes*), se trouve, au point de vue proprement morphologique, en infériorité décisive par rapport à la forme composée qui a pour tous les verbes une seule et même structure.

D'autre part la nuance de sens qui sépare fr. *j'ai aimé* de *j'aimai*, all. *ich habe geliebt* de *ich liebte*, vieux slave *neslŭ jesmi* « j'ai porté » de *nèsŭ* « je portai » etc., est souvent négligeable ; si le sujet parlant veut exprimer simplement le passé, il y parvient aussi bien par l'une que par l'autre forme. Il est ainsi conduit à recourir souvent à celle des deux formes dont la structure morphologique est le plus commode à manier. Peu à peu, il y a tendance à ne garder parmi les formes simples que celles qui sont fixées dans la mémoire, mais presque toutes sont des formes plus ou moins anomales, appartenant à des verbes forts, et que ne défend pas l'ensemble du type. On arrive ainsi à ne plus faire de distinction de sens entre la forme simple et la forme composée.

Dès lors, comme les langues ne gardent jamais deux formes grammaticales exactement synonymes, la forme difficile est éliminée au profit de la forme commode. Là où la distinction

de sens entre le parfait et le prétérit historique entre mal dans le plan de la langue, comme en germanique ou bien là où, comme en slave, il existe d'autres nuances de sens qui rejettent celle-ci dans l'ombre, l'élimination de la forme simple est assez aisée. Des circonstances de détail contribuent souvent à l'élimination. Mais ce qui domine le fait, ce sont les principes indiqués ci-dessus.

L'utilité des observations sur la morphologie générale est précisément qu'elles permettent d'apprécier dans une certaine mesure le degré d'importance des causes auxquelles on doit attribuer les innovations. On est amené à éliminer toutes les causes qui seraient particulières à une langue, dès l'instant qu'il s'agit d'un fait constaté sur un grand nombre de points : si une même évolution se produit sur deux domaines distincts, ce peut être dû à une rencontre fortuite, mais si on l'observe sur cinq ou six grands domaines, le hasard semble exclu et il faut découvrir des causes qui aient pu agir sur tous les domaines considérés. La généralité même du fait est une donnée de premier ordre pour la recherche des causes.

Remarque (du 5 mai 1920).

Je suis heureux de constater que, dans un des chapitres de sa suggestive brochure *La faillite de l'étymologie phonétique* (Neuveville [canton de Berne], 1919), M. J. Gilliéron est arrivé indépendamment à la même conclusion.

En son style imagé, M. Gilliéron écrit : « Avec elle (la disparition du passé défini), commence pour les verbes une nouvelle ère, c'est un acheminement vers l'état du verbe où il n'y aura plus comme voiles que des auxiliaires faisant manœuvrer une coque qui porte l'idée. »

M. Gilliéron insiste peu. Mais l'idée est capitale. Il s'agit d'une révolution dans la langue. Par l'élimination du prétérit simple, le français commence pour les verbes une évolution pareille à celle qui est à peu près achevée pour le substantif.

Le trait essentiel de la structure morphologique de l'indo-européen, et encore du latin, c'est que le mot n'existe pas indé-

pendamment de la forme grammaticale : il n'y a pas un mot signifiant « cheval », il y a un nominatif singulier *equus*, un génitif singulier *equī*, un accusatif pluriel *equōs*, etc., et l'on ne saurait isoler aucun élément signifiant « cheval » indépendamment des finales. Au contraire, dans le type moderne représenté par l'anglais, et, un peu moins bien par le français, le mot tend à exister indépendamment de tout « morphème » : quel que soit le rôle joué dans la phrase, on dit en anglais *dog* et en français *chien*, là où le latin avait une série de formes suivant les cas.

En ce qui concerne le substantif, le développement est achevé en français comme en anglais : une opposition comme celle entre *ʃəval* et *ʃəvo* (*cheval* et *chevaux*) n'est, dans le français actuel, qu'une survivance contre laquelle proteste le sentiment intime des sujets parlants, que les parlers locaux et le langage enfantin tendent à éliminer, et que, seul, maintient le conservatisme rigide du français normal : *cheval* tend à être aussi fixe que l'anglais *horse*.

Au contraire, en français, à la différence de l'anglais, le verbe est encore un mot-forme. Il n'y a pas en français un mot signifiant *finir*, mais un jeu de *je finis*, *nous finissons*, *je finirai*, *j'ai fini*, etc. Il n'y a surtout pas un mot signifiant *vouloir*, mais un jeu, très compliqué, de *je veux*, *nous voulons*, *je voudrai*, *j'ai voulu*, etc. Le français est encore bien loin de la simplicité de l'anglais, où un verbe *end* — identique au substantif *end* — se retrouve identique dans toutes les situations, pourvu seulement de quelques affixes, les mêmes qui figurent dans tous les verbes normaux, et déterminé par quelques pronoms et auxiliaires, ce mot presque invariable suffit à toutes les fonctions.

Mais le français tend à se rapprocher du type où l'anglais est déjà parvenu. La conjugaison du type *chanter* a déjà au présent quatre de ses formes sur six qui sont identiques *ʃɑ̃t* (*je chante*, *tu chantes*, *il chante*, *ils chantent*), et cette forme sert aussi pour le subjonctif et pour l'impératif singulier. Il suffit d'y ajouter quelque affixes, très réguliers, pour en tirer toutes les autres formes : *chantez*, *chantais*, *chanté*, *chanterai*, etc.

L'élimination du prétérit simple n'a pas seulement fait dispa-

raître des formes tout à fait aberrantes dans les verbes forts : *je fis*, à côté de *je fais, faire* ; *je sus*, à côté de *je sais, savoir* ; *je dois*, à côté de *je dois, devoir* ; etc. Elle a aussi supprimé l'une des flexions le plus compliquées du verbe français, et celle qui présentait les formes les plus éloignées de toutes les autres : *je chantai, tu chantas*, et *il chanta, nous chantâmes, vous chantâtes, ils chantèrent*. Sans doute le futur, *je chanterai*, et *vous chanterez, tu chanteras* et *il chantera, nous chanterons*, et *ils chanteront*, a aussi trois formes différentes (abstraction faite de l'orthographe) ; mais ces formes concordent pour la plupart avec celles du verbe auxiliaire très usuel, *j'ai*, et elles concordent aussi pour la plupart avec des formes qu'on retrouve au présent : *nous chantons, vous chantez*.

Située dans l'ensemble dont elle fait partie, l'élimination du prétérit simple est un moment du grand développement qui entraîne les langues indo-européennes à passer du mot-forme variable au mot fixé une fois pour toutes. *J'ai aimé, tu as aimé*, etc. sont plus près du terme de cette évolution que ne l'étaient *j'aimai, tu aimas*, etc.

Avec son sens si juste de la réalité linguistique, M. Gilliéron a indiqué d'un mot la portée de ce grand fait auquel les historiens du français n'ont pas jusqu'ici prêté l'attention qu'il mérite.

LE

RENOUVELLEMENT DES CONJONCTIONS[1]

Toutes les conditions qui déterminent la stabilité des mots semblent réunies pour assurer la persistance des conjonctions au cours du développement des langues.

La fréquence d'emploi fixe les mots dans la mémoire des sujets parlants. Les mots les plus employés sont du reste ceux que connaissent et dont font usage tous les membres de la communauté linguistique ; ce sont par suite ceux qui ne peuvent changer que par un consentement commun de tous les sujets, en quelque sorte, et qu'il est le plus malaisé de modifier. Or, certaines conjonctions sont d'un emploi constant. Et même les conjonctions moins usuelles reviennent dans des phrases multiples.

Les mots sont sujets à varier quand les choses exprimées viennent à changer ; or, les conjonctions expriment des notions permanentes dont rien ne provoque la variation. De plus, les mots sont souvent remplacés quand des usages sociaux, des croyances, des modes en provoquent l'élimination partielle ou totale, transitoire ou durable ; à toutes ces causes de changements les conjonctions échappent naturellement : il n'y a pas d'interdiction de vocabulaire qui puisse les atteindre.

Souvent le vocabulaire se modifie quand des mots sont empruntés à des langues étrangères et qu'on vient à substituer des termes étrangers, qui semblent plus élégants, à des termes de la langue indigène. Les conjonctions sont peu sujettes à

1. *Annuaire* de l'École pratique des Hautes Études, section historique et philologique, 1915-1916.

l'emprunt ; elles n'y échappent pas tout à fait, il est vrai, et le turc osmanli littéraire, par exemple, a emprunté au persan sa conjonction *ki* « que » ; mais c'est chose tout à fait exceptionnelle.

Enfin les conjonctions sont des particules dont le rôle est quasi grammatical : les grammaires leur font régulièrement une place. Or, on sait que la morphologie est ce qu'il y a de plus durable dans les langues. M. Terracher a montré, très justement, que des parlers locaux français dont le vocabulaire est tout pénétré d'éléments non locaux ont gardé dans leur grammaire beaucoup de traits anciens. Et c'est par la morphologie à peu près uniquement qu'on peut faire la classification généalogique des langues.

Les conjonctions sembleraient donc devoir se perpétuer d'une manière presque indéfinie, et les langues d'une même famille devraient présenter les mêmes conjonctions sans autre changement que ceux que nécessitent les variations de la prononciation.

Or, on observe en fait une situation exactement inverse de celle qui est attendue. Les diverses langues indo-européennes offrent des conjonctions très différentes les unes des autres ; quelques conjonctions, comme gr. ἵνα, sont d'origine complètement obscure, l'étymologie de beaucoup est mal connue. Les langues romanes ont aussi, dans une large mesure, des conjonctions différentes, et dont beaucoup ne continuent pas une conjonction latine : la plus importante des conjonctions françaises de subordination, *que*, ne représente pas exactement une conjonction du latin ancien ; le roumain, en particulier, a des conjonctions très différentes de celles de la plupart des autres langues romanes. De même, l'allemand et l'anglais sont loin de concorder : *und* et *and* ne sont pas identiques ; *wenn* et *if* sont tout à fait distincts. Même d'une langue slave à l'autre, les différences sont notables, et, dès qu'il s'agit de conjonctions, l'étroite ressemblance des langues slaves entre elles n'apparaît plus.

On connaît très peu de conjonctions qui soient sûrement de date indo-européenne. Et les deux seules conjonctions indo-

européennes dont l'existence est bien établie ont disparu au cours de l'histoire des langues où on les observe.

La conjonction signifiant « et » était l'enclitique qui est représentée par skr. *ca*, zend *ča*, v.-perse *čā*, gr. τε, lat. *que*, c'est-à-dire un indo-européen **k^{w}e*. Courante à date ancienne, cette particule n'a survécu ni dans les langues de l'Inde, ni en iranien, ni en grec moderne, ni dans les langues romanes. Le gotique a encore un reste de i.-e. **k^{w}e* dans le *-h* de *nih* « et ne pas » ; mais, à la longue, cette particule enclitique ne s'est pas maintenue en germanique, pas plus qu'en indo-iranien, en grec ou en latin ; et, en gotique même, *-h* n'existe déjà plus à l'état isolé.

La conjonction signifiant « ou » était aussi un petit mot enclitique, représenté par skr. *vā*, v.-perse *vā*, lat. *ue*. Ce petit mot se retrouve, élargi, dans le koutchéen (tokharien B) *wat*. Le grec ne l'a déjà plus à l'état isolé, et ne le connaît que groupé avec ἠ-, dans ἠ(*F*)ε, qui se trouve chez Homère, et qui a abouti à ἤ de l'attique. Cette particule indo-européenne **wĕ* s'est éliminée au cours de l'histoire des langues, tout comme **k^{w}e*. Déjà même le grec ancien ne l'offre plus à l'état isolé.

Si l'on ne possédait pas d'anciennes formes des langues indo-européennes, attestées avant l'époque chrétienne, on n'aurait aucune idée précise de ces deux particules qui ont tenu une si grande place en indo-européen. L'élimination assez précoce de i.-e. **k^{w}e* et de i.-e. **wĕ* peut tenir, et tient assurément, en grande partie à ce que ces particules enclitiques sont d'un caractère particulier à l'indo-européen, et ne pouvaient se maintenir après que le type indo-européen avait commencé à s'altérer profondément : les mots accessoires de la phrase postposés aux mots principaux sont quelque chose de spécifiquement indo-européen. Une cause spéciale, de caractère syntaxique, a donc pu contribuer ici à l'élimination des deux conjonctions indo-européennes les plus importantes.

Mais ce qui montre que la tendance universelle à éliminer les outils qui servent au groupement des phrases est intervenue dans la disparition de **k^{w}e* et de **wĕ*, c'est que le relatif indo-européen a eu le même sort. L'existence d'un relatif indo-européen est

établie par la concordance de skr. *yáḥ*, avestique *yō*, gr. ὅς, v.-sl. *ji-že* (le neutre *je* existe encore sans la particule *že*) ; le lituanien a connu *ja-*, qui subsiste largement dans des conjonctions (voir Hermann, *Ueber die Entwicklung der lit. Konjunktionalsätze*), et le gotique a aussi *jabai* « si », qu'on n'a pas de raison de séparer. L'extension considérable de ce relatif exclut l'idée qu'il aurait eu, en indo-européen, une existence seulement dialectale. Or, le latin, bien qu'assez anciennement attesté, ignore tout à fait le relatif **yo-*, aussi bien que le dialecte voisin, le celtique. Et, dans les diverses langues où il est connu, ce relatif s'est éliminé de bonne heure ; le slave ne le connaît que dans les plus anciens textes ; le grec moderne ne l'a plus. En iranien, l'avestique l'a encore, mais le vieux-perse n'en possède que des traces dans des conjonctions fixées (voir A. Meillet, *Grammaire du vieux-perse*, p. 174 et suiv.). Ce n'est que par hypothèse qu'on croit encore l'entrevoir en celtique (voir Pedersen, *Vergl. Gramm. der kelt. Sprachen*, II, § 545, 3, p. 234 et suiv.). Le vieux relatif s'est donc éliminé partout, sauf dans l'Inde, et sans raison particulière. Malgré toutes les raisons de persister qu'avait ce mot, il a fini par ne survivre presque nulle part, et c'est à peine si, dans quelques langues de type très archaïque, il en subsiste encore actuellement, hors des langues modernes de l'Inde, des restes méconnaissables.

Si l'on connaît très peu de conjonctions indo-européennes, ce n'est donc pas nécessairement parce que l'indo-européen ne liait pas les phrases, parce qu'il usait, comme l'on dit, de la « parataxe ». Sans doute, comme toutes les langues non écrites, l'indo-européen juxtaposait souvent les phrases, et la subordination n'y avait pas — à beaucoup près — l'importance qu'elle a prise dans les langues littéraires ; la langue parlée, qui dispose de toutes les variations de durée, de hauteur et d'intensité, de la suspension de la voix et, de plus, du geste, n'a pas besoin d'exprimer par des mots spéciaux les liens entre les idées autant que le fait la langue écrite. Mais l'absence totale de conjonctions de subordination ne se concilie pas avec l'existence d'un relatif tel que celui qui est attesté, et qui avait entièrement le caractère

d'un relatif proprement dit. La manière dont le relatif et des conjonctions, conservés d'abord, se sont éliminés au cours de l'histoire de certaines langues montre qu'il serait illégitime de conclure de l'absence de toute conjonction de subordination parmi les mots indo-européens établis à l'absence de tout élément de ce genre en indo-européen. Les conclusions *ex silentio* sont toujours dangereuses en grammaire comparée ; ici, elles conduiraient à une grossière erreur. Le fait que lat. *sed*, *nam*, et surtout *cum* et *ut*, ne sont pas représentés dans les langues romanes n'empêche pas ces mots d'avoir été très employés en latin, et dans le parler le plus courant, le plus familier. Nulle part plus qu'en matière de conjonctions, il ne faut se méfier de la conclusion que le comparatiste est toujours porté inconsciemment à tirer : puisqu'aucune comparaison ne permet de restituer un élément ancien, c'est qu'il n'y avait rien. Là où l'on peut suivre l'histoire des langues en fait, cette conclusion apparaît ruineuse, on vient de le voir.

Il reste à rechercher pourquoi, en dépit des circonstances qui paraissent de nature à en assurer la stabilité, les conjonctions et le relatif sont si sujets à disparaître et à se renouveler. Si l'on arrive à déterminer les causes de ces disparitions et de ce renouvellement, l'histoire du vocabulaire en sera sans doute éclairée d'une manière générale, et l'on pourra ainsi se rendre compte de la direction dans laquelle on a chance de trouver l'explication des conjonctions. On comprendra aussi pourquoi, dans les langues dont on ne possède pas toute l'histoire et où, seule, la comparaison permet d'expliquer l'origine des faits, l'étymologie de bien des conjonctions demeure nécessairement obscure.

La première et la plus importante de ces causes consiste dans le besoin qu'éprouve le sujet parlant d'être expressif, de bien faire sentir sa pensée et d'agir sur son interlocuteur.

Le fréquent emploi d'un mot, par le fait même qu'il l'imprime dans la mémoire et en rend l'usage pour ainsi dire automatique, lui enlève toute valeur expressive. Or, cette perte de la valeur expressive, si elle fournit des outils abstraits com-

modes pour la langue écrite, a le plus grand inconvénient pour la langue parlée, qui ne peut exister sans un appel constant à l'attention et à la sensibilité de l'interlocuteur. Des conjonctions françaises comme *et*, *ou*, *que*, *si*, etc., sont d'usage courant. Mais elles n'ont aucune valeur sentimentale ; ce sont de simples accessoires de phrases, et leur sens est très abstrait. Il en est de même des conjonctions de même sens dans les diverses langues littéraires de l'Europe moderne : fr. *si*, all. *wenn*, angl. *if*, russe если, etc. ne sont que les signes de la phrase conditionnelle et n'ont pas plus de valeur expressive que n'en a une forme grammaticale.

Quand on parle, on est naturellement conduit à réagir contre cette absence totale d'expression à quoi viennent aboutir de manière nécessaire les conjonctions. Ceci se manifeste bien en français où, par exemple, dans la langue parlée, *et* tend à être remplacé par *et puis*, *et après*, *et alors*, *et de plus*, ou même, à Paris, dans le parler populaire, par *et puis alors* (*et puis* étant déjà très affaibli), dans la jonction de deux phrases, tandis que *et* suffit encore bien à relier deux mots à l'intérieur d'une phrase. De même fr. *ou* reliant deux phrases tend à s'élargir en *ou bien* : *ou* seul manque de force ; l'italien a *ossia*, etc.

L'affaiblissement de la valeur expressive des conjonctions ne tient pas seulement à l'usure qui résulte de la fréquence de leur emploi. Il s'y joint une diminution de ce que l'on pourrait appeler leur « volume phonétique ». Les événements phonétiques de l'histoire des langues ne tendent presque jamais qu'à une diminution de l'effort, et, par suite, à une réduction des mots existants. Un mot qui se perpétue dans une langue en subissant seulement tous les changements de prononciation successifs, sans aucune autre altération, se réduit peu à peu ; ainsi une troisième personne telle que **legeti* se réduit à lat. *legit* et à fr. *lit* (prononcé *li* simplement). Ceci entraîne souvent des élargissements par réaction ; ainsi un mot comme l'indo-européen **owis* « mouton », encore exactement conservé en latin sous la forme *ouis*, aurait donné en français un monosyllabe, réduit à une seule voyelle ; ce mot a été élargi dès le latin « vulgaire » par un

suffixe assez lourd, et l'on a en français *ouaille* (*oveille* en vieux français) ; le mot *ouis* a même disparu dans la plupart des langues romanes, comme οἷς en grec.

La conjonction *et* du français, dont le *t* n'est que graphique et ne se prononce jamais même devant voyelle, illustre bien ce fait : en latin, c'était *et*, avec un *t* constamment prononcé, et l'italien a encore *ed* devant voyelle. La forme it. *e*, fr. *é* (suivant la prononciation) est donc réduite de moitié par rapport à la forme latine, et cette réduction de « volume phonétique » contribue largement à la diminution de « valeur expressive ». En latin même, *et* représente un plus ancien **eti*, mot qui, originairement, ne signifiait pas « et », mais « de plus, encore », et qui correspondait à gr. ἔτι « encore », à skr. *áti* « en outre ». La réduction progressive dont fr. *et* (prononcé *é*), lat. *et*, i.-e. **éti* donnent une idée si nette n'est pas exceptionnelle : en arménien, par exemple, on voit comment l'indo-européen **épi* « de plus, aussi », attesté par skr. *ápi* « en outre, aussi », gr. ἔπι « dessus », donne la forme arménienne ancienne *ew*, où tout est conforme au développement phonétique normal : chute de la voyelle de syllabe finale et passage de *p* intervocalique à *w*. L'arménien ancien *ew* a abouti à la simple voyelle *u* dès le haut moyen âge, si bien que l'i.-e. **epi* est actuellement représenté en arménien par la seule voyelle *u*.

Ce fait arménien fait apparaître une autre circonstance qui a souvent contribué à diminuer le « volume phonétique » des conjonctions. Les conjonctions ne sont pas dans les phrases des mots principaux ; ce sont des mots *accessoires*. Or, les mots accessoires tendent à se prononcer plus vite que les mots principaux, et leurs éléments constituants, plus réduits par là même, sont plus sujets à s'altérer. C'est ce qui est arrivé quand l'ancien arménien *ew* a passé à *u*.

Ainsi, tantôt par voie sémantique, tantôt par voie phonétique, et tantôt par une combinaison des deux ordres de conditions, les conjonctions tendent constamment à s'affaiblir et à se réduire.

C'est avant tout le besoin de renouveler l'expression pour en

raviver la force qui a conduit à remplacer l'ancien relatif indo-européen par les formes de l'interrogatif.

Le besoin de renforcer le vieux relatif indo-européen est déjà très sensible en grec ancien, où le relatif est élargi par la particule περ dans bien des cas, et où l'on trouve ὅσπερ ; en slave, la chose est plus manifeste encore, puisque le relatif *je-* ne s'emploie en tant que relatif qu'avec la particule *že*, et qu'on a toujours v.-sl. *jiže* (ancien *ji-že*) au nominatif masculin singulier. Mais ce rafraîchissement du relatif par l'addition d'une particule n'a suffi nulle part à la longue. On a été amené à recourir à des mots nouveaux.

Le procédé le plus simple de renouvellement a consisté à utiliser le démonstratif : le démonstratif **to-* servait en indo-européen à rappeler un mot d'une phrase précédente ; on a tiré parti de cette propriété en allemand où *der* a pris, par sa situation à la limite de deux phrases, la valeur relative ; le même phénomène s'était produit beaucoup plus tôt en perse, où, déjà dans les inscriptions achéménides, au VI^e^ siècle av. J.-C., la seule forme connue du relatif est le démonstratif *hya, tya-*, et où il ne subsiste de souvenir de l'ancien relatif indo-iranien *ya-* que dans des conjonctions comme *yaθā* « comme ». Le gotique a de même utilisé un démonstratif anaphorique dans son type *izei*. Ce procédé fournissait très commodément un relatif nouveau, mais le relatif ainsi obtenu n'avait rien de particulièrement expressif. Aussi n'est-il pas le plus ordinaire.

Le procédé le plus fréquent dans les langues indo-européennes, à savoir le passage de l'interrogatif à la valeur relative, se laisse bien illustrer au moyen de phrases de l'évangile en grec. La valeur interrogative de τίς est encore très nette dans la phrase, Luc, VI, 11, διελάλουν πρὸς ἀλλήλους τί ἂν ποιήσαιεν τῷ Ἰησοῦ. Mais le second membre de cette phrase est une subordonnée, comme on le voit par l'optatif qui y figure (et qui, dans le Nouveau Testament, n'est du reste qu'une survivance : la langue de Luc est relativement littéraire, on le sait) ; la traduction en vieux-slave *glagolaaxǫ drugŭ kŭ drugu čito ubo bišę sŭtvorili Isusovi* (d'après le Zographensis), avec le « conditionnel » slave, présente les mêmes

caractères que le texte grec ; de même la traduction arménienne *xawsein ənd mimeans the zinč aṙniçen Yisusi* ; de même la traduction gotique *rodidedun du sis misso ƕa tawidedeina* (*tawidideina* ms.) *þamma Iesua*. Cette phrase peut se traduire par « ils s'entretenaient de ce qu'ils pourraient faire à Jésus », traduction qui fait ressortir la valeur presque relative de τί et de ses équivalents slave, arménien et gotique.

Les phrases de cette sorte ont pu servir de modèles pour d'autres où la valeur interrogative n'est plus qu'à peine appréciable ; ainsi : Luc, VIII, 36, ἀπήγγειλαν αὐτοῖς οἱ ἰδόντες πῶς ἐσώθη ὁ δαιμονισθείς. La langue du Nouveau Testament n'est pas allée plus loin ; l'ancien interrogatif n'y sert encore de relatif que dans la mesure où des phrases à τίς ayant valeur interrogative pouvaient servir de modèles. En dehors de ces cas, le relatif est toujours ὅς. Mais, avec le temps, l'emploi de l'interrogatif avec valeur relative s'est étendu, et c'est la particule invariable που qui, en grec moderne, est l'unique mot relatif. On observe des faits analogues ailleurs ; car la traduction slave de la phrase citée est : *vŭzvěstišę imŭ viděvŭšei i kako sŭpase sę běsŭnovavy* (d'après le Zographensis et le Marianus) ; la traduction arménienne , *patmeçin noça oroç tesealn ēr the ziard phrkeçaw diwaharn* ; la traduction gotique : *gataihun þan im jah þai gasaihwandans hwaiwa ganas sa daimonareis*.

Le procès qui vient d'être décrit n'est peut-être pas le seul par lequel l'interrogatif ait pris la valeur relative ; on en a imaginé d'autres (voir Delbrück, *Vergleichende Syntax*, III, p. 389 et suiv.) ; c'est presque évidemment le principal ; et c'est le seul qu'on puisse suivre en fait dans la période historique de certaines langues.

Tous les procès imaginables reviennent à ceci que la langue a remplacé un tour abstrait et inexpressif par un tour qui, au moment où a eu lieu la substitution, était expressif. On peut dire que l'interrogatif reçoit la valeur relative au moment où, d'après le modèle de phrases ambiguës comme celles citées ici dans lesquelles l'interrogatif pouvait être conçu comme relatif, ce même interrogatif vient à être employé dans des phrases où il n'y a aucun sens interrogatif.

Or, la substitution de l'interrogatif au relatif n'est pas chose rare : elle a eu lieu dans tous les dialectes slaves, postérieurement à l'époque slave commune, puisque *iže* est encore le seul relatif du vieux-slave ; on l'observe en lituanien, dans les dialectes germaniques, en latin et dans les autres dialectes italiques, en « tokharien », en albanais, en arménien (où le relatif *or* est identique à l'adjectif interrogatif *or*), en perse ; le fait perse est particulièrement remarquable parce que, après la période du vieux-perse, où le relatif *ya-* a été remplacé par un démonstratif, est venue la période pehlvie, où ce démonstratif devenu relatif a été remplacé par l'ancien interrogatif (*Mémoires de la Société de linguistique*, XVIII, p. 242 et suiv.). Le fait italique est très ancien et paraît remonter à la période italique commune (au contraire, le celtique suit ses voies propres) ; car la différenciation de l'interrogatif lat. *quis*, osq. *pis*, ombr. *pis* et du relatif lat. *quī* (de **quoi*), osq. *pui*, ombr. *poi* est italique commune ; dans d'autres langues la répartition de **kʷo-* et de **kʷi-* est tout autre ; ainsi en slave où l'un de ces thèmes sert pour les personnes : v.-sl. *kŭto* « qui », et l'autre pour les choses : v.-sl. *čĭto* « quoi » ; et de même en arménien. Un développement qui s'est produit indépendamment, à des dates différentes, dans tant de langues si diverses répondait évidemment à une tendance naturelle. Dans le passage du latin au français, l'interrogatif *quare*, dont l'explication est évidente en latin et qui n'était qu'un juxtaposé, a encore donné la conjonction de coordination *car*, qui, après une période de vitalité, est devenue la plus abstraite des conjonctions et qui tend par suite à sortir de l'usage dans la langue parlée ; le latin *quare* n'est du reste pas représenté dans les langues romanes en dehors du groupe français et provençal.

Un point important, qu'il ne faut jamais perdre de vue, c'est que, dans toutes les langues où l'on observe le passage de mots au rôle de conjonction ou de relatif, il existait déjà des conjonctions et un relatif et que, par suite, ces mots n'ont eu qu'à se conformer à des modèles existants. On se représente trop souvent ce passage comme une véritable création à expliquer de toutes pièces ; en réalité, il ne s'agit jamais que de renouvellements ;

Brugmann a, très justement, insisté sur ce point *Indogermanische Forschungen*, IV, p. 229 et suiv., en ce qui concerne le relatif. Même si relatif et conjonctions sont d'un usage peu fréquent, il suffit que la catégorie grammaticale du relatif et des conjonctions existe pour que le développement d'un relatif nouveau et de conjonctions nouvelles en soit facilité d'une manière essentielle.

Les origines des conjonctions sont d'une diversité infinie, on le sait. Il n'y a pas d'espèce de mot qui ne puisse livrer des conjonctions. Le verbe même en fournit, comme lat. *uel*, *igitur* ou *licet*.

Quel que soit le point de départ, le trait commun à tous ces développements consiste en ce que, par l'effet de la répétition qui en a atténué progressivement la valeur expressive et en a fait oublier la signification propre, l'élément qui figure à la jonction de deux phrases tend à devenir un simple outil grammatical : il se « grammaticalise » pour ainsi dire. Le sens initial de l'élément devient chose à peu près négligeable et sans conséquence pour le développement ultérieur. L'exemple typique de ces « grammaticalisations » résultant de l'emploi de certains mots dans la phrase est le cas bien connu des mots français adjoints aux phrases négatives, *pas*, *point*, *rien*, *personne*, *jamais*, mots dont aucun n'avait par lui-même le caractère négatif, mais qui, par suite de la façon dont ils ont été employés dans la phrase, sont devenus des particules à valeur négative pleine, suffisant entièrement par elles-mêmes à indiquer la négation, et qui, dans la langue parlée, tendent à éliminer tout à fait l'ancienne négation *ne*, devenue désormais superflue. Cet exemple montre assez à lui seul que, dans l'étymologie des particules et mots accessoires de phrases, le sens initial des mots considérés est ce qui importe le moins : il n'y a rien de commun entre les sens initiaux — du reste très différents — de *pas*, *point*, *rien*, *personne*, *jamais* et la négation. Tous ces mots n'en sont pas moins pleinement négatifs aujourd'hui. Entre *pas* négatif et le substantif *pas*, il n'y a plus rien de commun ; *personne* négatif est maintenant un autre mot que le

substantif *personne* ; la valeur non négative de *jamais* ne subsiste guère que dans une langue artificiellement archaïsante ; elle n'est plus de l'usage parlé actuel.

On voit à quel point il serait vain de chercher dans le sens initial du mot qui a fourni une conjonction le principe du développement de sens de cette conjonction ; c'est le rôle dans la phrase qui décide de tout. On a essayé de classer les divers procédés par lesquels un adverbe devient ainsi conjonction (voir par exemple Morris, *On principles and methods of latin syntax*, p. 171 et suiv.). En réalité, il y a autant de procès distincts que de mots, et chaque cas est à étudier séparément.

Soit un mot comme le slave *da*. Cette particule est propre au slave et n'a de correspondant exact dans aucune autre langue indo-européenne. Mais le seul examen des faits slaves suffit à montrer que, entre les emplois principaux qu'on observe à l'époque historique et le sens originaire de *da*, il n'y a plus de lien actuellement visible. Le sens premier était sans doute « ainsi », et c'est par là que s'explique le sens de « oui » qui est celui de la particule à l'état isolé en russe, en bulgare, en serbo-croate — mais qui n'est pas attesté en vieux-slave, car gr. ναί y est traduit par *ei* et non par *da*. Un mot signifiant « ainsi », mis devant un verbe qui exprime une exhortation, un ordre, tend à prendre la valeur d'une particule d'exhortation ; et c'est en effet une valeur courante de *da* en vieux-slave. Mais si l'on juxtapose deux phrases dont l'une comporte le *da* exhortatif, et qu'on relie ces deux phrases par la prononciation, *da* fait l'effet d'un mot qui exprime le but de l'action ; *da* est ainsi arrivé à introduire une phrase finale et à prendre la valeur de « pour que, afin que », qui est la plus fréquente en vieux-slave. Un mot signifiant « ainsi » est tout aussi propre à introduire une phrase qui indique une condition et à prendre, par ce procédé, le sens de « si » : c'est ce que l'on observe notamment en bulgare, en vieux-serbe, en vieux-russe. Il se trouve ainsi qu'un même mot a pu exprimer la notion de « oui » et celle de « si » ; le même fait se rencontre dans le domaine latin, où *sī* est un mot ayant signifié « ainsi » et où *sīc* « ainsi », qui est *sī* plus une particule

c(e), a pris le sens de « oui » (ital. *si*). Quant au fait qu'un même mot sert à indiquer « pour que » et « si », il n'a rien non plus de surprenant en lui-même ; l'arménien *the* « que » et « si » en fournit du reste l'équivalent ; le sens initial de arm. *the* comme de sl. *da* est « ainsi ».

Les conjonctions peuvent ainsi se renouveler très aisément, puisque toute particule, ou même tout mot employé comme accessoire de phrase, tend à perdre son sens propre pour prendre une valeur due simplement à son rôle dans certaines phrases. Au moment où a lieu le renouvellement, on obtient une expression relativement fraîche et intense ; dès que le procès est terminé, il ne reste plus rien de cette force qui est chose toute transitoire. L'histoire des conjonctions se ramène presque tout entière à un effort toujours répété et, par nature, perpétuellement inutile pour obtenir des tours de phrase expressifs.

La conjonction finale du latin était *ut*, que le sens initial « comme » ne désignait pas particulièrement pour cet emploi. Devenu trop faible pour ce sens qui exige une expression assez forte, *ut* a disparu dans toutes les langues romanes. Le fr. *que*, qui en a pris la place, n'existe plus avec cette valeur que dans le tour : *Ote-toi que je m'y mette*, et la conjonction finale du français est *pour que* (ou, surtout dans la langue écrite, *afin que*). Le fr. *pour que* est trop faible actuellement et paraît insuffisant.

Il va sans dire que tous les types d'expression auxquels servent les conjonctions ne tendent pas également à se renouveler. La simple liaison, à laquelle sert la conjonction *et*, n'a pas besoin en général d'être exprimée d'une manière bien intense, et la conjonction qui sert pour « et » peut se maintenir longtemps sans être remplacée. L'expression de « ou » est déjà plus sujette à varier : outre *ue* qu'il conserve, le latin s'est donné *aut*, *uel*, *sīue* (*seu*) ; les langues romanes n'ont gardé que l'une, à savoir, *aut* presque partout, *sīue* en roumain ; le français emploie *soit... soit*. D'une manière générale, les conjonctions indispensables à l'usage le plus courant de la langue et qui ont presque le caractère de particules grammaticales, comme fr. *et*, *ou*, *que*, *si*, participent assez largement à la stabilité qui caractérise les éléments

grammaticaux du langage ; elles sont très anciennes en français ; elles remontent à la forme du latin sur laquelle reposent les langues romanes et ont leurs correspondants dans la plupart des autres langues de ce groupe. Au contraire les conjonctions à sens plus spécial et qui ont plus d'autonomie dans la phrase sont plus instables.

L'opposition, par exemple, doit être signalée à l'attention : si l'on indique un contraste entre deux phrases, c'est d'ordinaire pour y insister, et les conjonctions qui ont la valeur de « mais », d'une part, sont souvent variées, d'autre part, sont sujettes à être remplacées par d'autres. Les conjonctions qui signifient « mais » sont multiples en latin ancien : *at*, *ast*, *sed*, *uerum*, *uero*, *autem* expriment ainsi des degrés divers et des nuances variées de l'opposition. Aucune de ces conjonctions n'a survécu dans les langues romanes, où presque partout l'idée de « mais » est rendue par un mot nouveau, dont le sens ancien était « plus », lat. *magis* (it. *ma*, fr. *mais*, esp. *mas*), procédé qui se retrouve, par exemple, dans le serbe *vȅ* (ancien *vęće*, littéralement « plus grandement ») ; après une phrase négative, il y a une autre conjonction, également nouvelle, issue d'un mot signifiant « plus tôt » : it. *anzi*, v.-fr. *ainz*, esp. *antes*. Ces expressions de l'opposition n'ont pas suffi, et l'on a créé, par exemple, l'expression : it. *tuttavia*, v.-fr. *toutes voies*, esp. *todavia* ; le français a de nouveau rafraîchi *toutes voies* en *toutefois*, qui actuellement sort de l'usage parlé, comme *ainz* en est déjà sorti. Le fr. *mais*, qui est maintenant la seule conjonction adversative employée dans la langue courante, est devenu très débile et ne suffit plus. On est amené à dire *mais bien*, *mais plutôt*, *mais au contraire*, *mais assurément* (*mais bien sûr* en langage familier), etc.

Le latin a de même toute une variété de mots pour rendre la notion de « quoique » : *cum* et *si*, qui ont des valeurs plus générales, y servent souvent ; et, en outre, avec une valeur spéciale, *etsi*, ou, avec plus de force encore, *etiam si*, *tamen etsi* (*tametsi*), ou encore *quanquam* et *quamuis*. Tout cela n'a pas encore suffi, et, à l'époque impériale, la langue littéraire admet *licet*, ce qui montre que l'emploi de *licet* comme conjonction s'était déjà

imposé dans l'usage courant. De ces diverses conjonctions du latin ancien, les langues romanes n'ont rien gardé. La concession peut s'exprimer par la simple juxtaposition de deux phrases ; un ton de voix approprié indique assez l'opposition. Quand la concession est exprimée, c'est par des tours nouveaux : it. *ancor chè, ben chè, giacchè*, v.-fr. *encore que, bien que, ja soit que*, etc. Le français moderne, qui a conservé *bien que*, recourt à une autre expression, que n'avait pas le vieux-français : *quoique*. La langue familière ne se contente pas de ce tour abstrait : elle a *quand même, quand bien même* (en langage populaire, *quand même que*) et surtout *malgré que*. La création est donc constante ici, par suite du besoin qu'on a d'exprimer la concession avec une force toujours renouvelée.

Les mots nouveaux qui rendent une même notion, tout en étant absolument divers, ont souvent au fond le même sens originel : latin *et*, arménien *ew*, slave *i*, grec καί n'ont rien à faire étymologiquement les uns avec les autres ; mais le sens de « aussi » y transparaît encore d'une manière évidente dans les formes anciennes de chacune de ces langues et n'a tout à fait disparu qu'avec le temps ; ce sens est encore très visible en latin classique, en arménien classique, en vieux-slave, en grec ancien.

Mais il arrive aussi que des conjonctions de même valeur aient des sens originels tout à fait divers. Ainsi pour « si », l'indo-iranien a encore une conjonction de la famille du relatif, skr. *yádi*, v.-perse *yadiy* (sens initial « quand ») ; le grec a une ancienne particule servant à l'exhortation, ion.-att. εἰ, dor. αἰ (cf. εἴ-θε, αἴ-θε) ; le latin et l'osco-ombrien des particules (du reste de formes distinctes en latin, d'une part, en osco-ombrien, de l'autre) signifiant « ainsi », lat. *sī* et osq. *svaí, suae*, ombr. *sve* ; etc. Dans la langue parlée, la condition s'exprime souvent par une simple juxtaposition ; on connaît, par exemple, le tour français populaire : *une supposition, il vient, je le chasse*.

Un détail est important à noter, qu'on a déjà signalé en passant : dans la conversation courante, dans la langue familière, où les inflexions de la voix et les pauses indiquent assez les

rapports entre les idées exprimées, on a relativement peu besoin de conjonctions ; on parle d'ordinaire par phrases simples, qui ne sont pas liées grammaticalement les unes aux autres.

Les conjonctions sont surtout utiles dans le discours solennel ou rituel, où l'on dispose beaucoup moins librement des inflexions de la voix et des pauses ; elles sont indispensables dans la langue écrite qui, sans conjonctions, devient aisément inintelligible. De plus, le langage solennel et la langue écrite recourent volontiers à des phrases composées de plusieurs membres pour exprimer une pensée complexe et nuancée. Tout en se développant le plus souvent dans le parler courant et expressif, les conjonctions se fixent donc surtout dans le parler solennel, et en particulier dans la langue écrite. Mais du parler familier, où elles servent dans leurs débuts, quand elles sont encore à peine des conjonctions, à insister sur les intentions du sujet parlant, elles passent à la langue savante, qui en développe l'usage et qui les conserve souvent assez longtemps, alors même qu'elles ont disparu de la langue parlée.

C'est cet usage savant des conjonctions qui fait que quelquefois on les emprunte à une langue étrangère : le turc a emprunté le *ki* persan. L'histoire des langues indo-européennes offre peu d'équivalents de ce fait. Toutefois le roumain a aussi emprunté des conjonctions. Il ne faut pas oublier que la langue savante, où les conjonctions sont le plus nécessaires, est particulièrement sujette à emprunter des mots de toutes sortes.

Si les éléments de liaison des phrases, et en particulier les conjonctions de subordination, diffèrent pour la plupart d'une langue indo-européenne à l'autre, cela ne tient pas nécessairement, on le voit, à ce que l'indo-européen aurait été une langue peu civilisée, pratiquant seulement la juxtaposition des phrases. De l'indo-européen on ne connaît guère que ce qui a subsisté des formes de l'ancienne langue courante dans plusieurs idiomes à la fois. Or, dans la langue courante, l'expérience montre que, d'une part, sauf quelques particules indispensables comme *et*, *ou*, *que*, *si*, les conjonctions sont relativement peu employées, et que, d'autre part, elles sont sujettes à se renouveler sans cesse.

SUR LES CARACTÈRES DU VERBE[1]

Les anciens grammairiens distinguaient une dizaine de « parties du discours ». Comme les principes de leurs distinctions n'étaient pas homogènes, et que d'ailleurs la situation diffère d'une langue à l'autre, ils n'arrivaient pas à fixer un nombre bien arrêté de ces catégories qui pourtant étaient censées représenter les éléments fondamentaux du langage.

Il n'y a, en réalité, que deux espèces de mots dont la distinction soit essentielle, commune à toutes les langues, et qui s'opposent nettement l'une à l'autre : la catégorie du *nom* et celle du *verbe*. Le nom indique les « choses », qu'il s'agisse d'objets concrets ou de notions abstraites, d'êtres réels ou d'espèces : *Pierre*, *table*, *vert*, *verdeur*, *bonté*, *cheval* sont également des noms. Le verbe indique les « procès », qu'il s'agisse d'actions, d'états ou de passages d'un état à un autre : *il marche*, *il dort*, *il brille*, *il bleuit* sont également des verbes. (On laisse de côté, à dessein, les mots accessoires, articles, pronoms, prépositions, conjonctions, etc., qui posent des problèmes particuliers.)

La distinction du nom et du verbe s'exprime toujours par quelque procédé grammatical. Elle n'est pas également marquée partout, à beaucoup près, et la forme en varie d'une manière très considérable suivant les langues.

Les langues où la distinction est le plus immédiatement visible sont les langues dites flexionnelles où les mots portent dans leur forme même la marque du rôle qu'ils jouent dans la phrase. Dans

1. *Revue philosophique*, t. LXXXIX (janvier-février 1920), p. 1 et suiv.

une langue comme le latin, on voit du premier coup que *ago*, *agis*, etc. est un verbe et que *dominus*, *domini*, etc. est un nom. Telle est la situation pour toutes les anciennes langues indo-européennes ou sémitiques : la « déclinaison » du nom s'y oppose à la « conjugaison » du verbe. Et ce n'est pas chose particulière à ces langues : le finno-ougrien, le caucasique méridional, le bantou offrent, chacun à leur manière, des faits analogues. Partout où il y a flexion, la flexion nominale et la flexion verbale se distinguent en s'opposant l'une à l'autre. Du reste, l'une des deux flexions peut disparaître sans que l'autre en souffre : beaucoup de langues indo-européennes, ainsi les parlers romans, l'anglais, le persan, ont perdu la déclinaison nominale, tout en gardant, plus ou moins complètement, la conjugaison verbale. De même, en sémitique, la déclinaison nominale a disparu, à des dates diverses, suivant les langues, tandis que la conjugaison verbale s'est largement maintenue partout.

Mais, même en l'absence de toute flexion, la distinction du nom et du verbe subsiste, exprimée par des moyens linguistiques. Le fait de placer un complément avant ou après un mot suffit en chinois à indiquer si ce mot est nom ou verbe. En anglais, un même mot *love* peut servir aussi bien de verbe que de nom; mais le substantif *love* « amour » ne se confond pas pour cela avec le verbe *love* « aimer » : s'il s'agit du nom, on dira *the love*, *a love*; s'il s'agit du verbe *I love*, *you love* ou, à l'infinitif, *to love* : le mot *love* reste le même, mais les mots accessoires qui l'accompagnent, et qui sont un des éléments essentiels de la morphologie anglaise, diffèrent. Et l'on dira *the man's love* « l'amour de l'homme », mais *I love the man* « j'aime l'homme ». Sans doute, il peut y avoir tel cas particulier où, en l'absence de toute particule et de tout complément, une forme *love* est ambiguë; mais c'est chose exceptionnelle; et l'ensemble des emplois donne aux sujets parlants le sentiment net qu'il y a un substantif *love* et un verbe *love*. Il ne faut pas envisager le mot isolé : ce n'est qu'une abstraction vaine; le mot ne se manifeste que dans la phrase, et, s'il est en règle générale employé d'une manière telle qu'on ne puisse se méprendre sur son rôle ou verbal ou nominal, on peut dire que

le mot est soit verbe, soit nom, même si rien dans sa forme propre ne le donne comme verbe ou comme nom, et s'il peut jouer l'un ou l'autre rôle suivant les cas.

Ces considérations autorisent à poser la catégorie du verbe comme essentielle dans le langage.

*
* *

Par le fait que le verbe exprime un procès, la langue est amenée à marquer à qui ou à quoi s'applique ce procès. Là où il y a flexion, le verbe exprime avant tout ce qui agit ou qui est dans un état donné ou qui passe d'un état à un autre : la flexion personnelle est la plus naturelle au verbe, et la plus ordinaire. C'est aussi celle qui persiste le plus. Même celles des langues indo-européennes qui ont le plus altéré les fins de mots et le plus réduit tout l'ensemble de la flexion ont gardé quelques traces de la flexion personnelle. En français, c'est avant tout le pronom — devenu un simple outil grammatical — qui marque la personne; mais, même dans une conjugaison comme celle d'*aimer* où, au présent, quatre des formes se sont confondues dans la prononciation (*j'aime, tu aimes, il aime, ils aiment*), il reste deux formes pourvues de flexion : *nous aimons, vous aimez*. En anglais, où *I love, we love, you love, they love* ont une même forme de *love*, il y a une forme propre pour *he loves*. Quand il a une flexion propre, le verbe se reconnaît donc le plus souvent à ce qu'il possède une flexion personnelle.

On voit bien en français, par l'affaiblissement de sens qu'entraîne l'habitude de les employer, comment des pronoms autrefois autonomes deviennent peu à peu de simples éléments grammaticaux, équivalant à des désinences : *je, tu, il, ils* ont été d'abord des pronoms et ont eu le caractère de mots autonomes; aujourd'hui, ce ne sont plus que des parties de formes verbales, dénuées de sens par elles-mêmes : un tour tel que : *je, huissier près le tribunal de X, signifie*, n'est qu'un archaïsme de la langue judiciaire ou administrative, et un archaïsme qui surprend un Français du temps présent. Les marques de la personne sont si nécessaires qu'elles

tendent à figurer même si la personne est exprimée déjà par un pronom expressif et vraiment autonome : *moi, je dis ; toi et moi, nous croyons*. Et même, dans le parler populaire, on tend à mettre *il, elle* devant le verbe à la troisième personne quand le sujet est un substantif : *la vache, elle mange ; vos amis, ils sont arrivés ; ton eau, elle bout* ; etc. Mais les formes du type du français : *j'aime, tu aimes, il aime, ils aiment*, montrent aussi que le pronom marquant la personne près du verbe est — ou du moins peut être — autre chose que le complément d'un nom. En sémitique et en hamitique, les affixes qui indiquent les personnes près des verbes sont manifestement identiques aux pronoms personnels ; l'exemple du français montre que ces pronoms ne doivent pas nécessairement être considérés comme l'équivalent de compléments de nom, comme l'ont parfois imaginé des linguistes qui croyaient à la priorité des formes nominales ; et, par exemple, une forme sémitique qui s'analyse *aime-moi* ne remonte pas nécessairement à un ancien type *aimer de moi*, où l'élément *aime* aurait une valeur nominale, comme on l'a souvent supposé : le pronom qui est devenu une marque de la personne peut avoir, comme en français, un tout autre caractère, et le radical peut avoir une valeur vraiment verbale. Du reste, en indo-européen où le verbe est distinct du nom plus qu'il ne l'est nulle part ailleurs, des marques de personnes employées dans la flexion verbale, les unes sont tout à fait distinctes du pronom personnel, les autres ne les rappellent que de loin, et l'on n'a par suite pas le droit d'identifier à des pronoms indo-européens telle ou telle des désinences de la flexion personnelle qui se trouve, peut-être par hasard, leur ressembler.

Et, en effet, la flexion personnelle, si caractéristique qu'elle soit du verbe, ne lui est pas essentielle. Elle manque dans beaucoup de langues, notamment dans les langues d'Extrême-Orient où il n'y a aucune flexion, et où, de plus, la personne s'exprime non par des pronoms, mais par des formules polies, du type de *sa grandeur, son excellence*.

Même dans les langues où la flexion personnelle est le plus développée, il est commode, pour construire les phrases, d'avoir

des formes capables d'exprimer le procès, mais qui admettent les mêmes constructions que le nom. Il s'est donc constitué des formes qui, tout en indiquant le procès, tout en étant des verbes, puisqu'elles admettent les mêmes compléments et les mêmes déterminations que les verbes, ont le caractère nominal ; on appelle infinitifs les formes de ce genre qui ont le rôle de substantifs, et participes celles qui ont le rôle d'adjectifs. Le nom de *participe* est, comme la plupart des termes grammaticaux, un mot latin, calqué sur un mot grec qui a servi de modèle, et qui signifie : « qui a part (à la nature du nom et du verbe tout à la fois) » ; le nom s'appliquerait à l'infinitif aussi bien qu'au participe, mais, par suite de la structure de la langue grecque où le mot a été formé et où l'infinitif n'était pas un vrai substantif, au moins par la forme, tandis que le participe était un véritable adjectif à tous égards, les Grecs qui ont formé le mot sur lequel *participe* est calqué ne l'ont appliqué qu'à des formes ayant valeur d'adjectifs ; on est ici en présence d'un de ces accidents dont est faite toute l'histoire du vocabulaire. Des formes comme les infinitifs latins *dicere* ou *dixisse* ou comme les participes *dicens*, *dicendus*, *dicturus*, *dictus* sont aussi verbales, en un sens, que *dico* ou *dixi*. Mais, de plus, elles admettent des constructions nominales, et elles n'admettent pas certaines constructions verbales : en général, elles ne servent que d'éléments secondaires, souvent en apposition, dans des phrases où figurent des formes verbales personnelles.

D'ailleurs le verbe n'est pas un élément constant de toute phrase. Il y a deux sortes de phrases : les phrases *nominales* où une chose est affirmée d'une autre chose, et les phrases verbales où est exprimé un procès. Dans les phrases nominales telles que : *l'homme est bon*, *la maison est neuve*, *il est chez toi*, il ne figure un verbe que pour la commodité de la phrase ; les langues sémitiques n'y ont aucune forme verbale ; l'indo-européen n'en avait pas le plus souvent ; le russe dit encore aujourd'hui *dom nov* « la maison est neuve », *on u tebê* « il est chez toi ». C'est dire à quel point l'analyse de *j'aime* en *je suis aimant*, où s'est complue longtemps la grammaire dite logique, est artificielle, et loin de la réa-

lité. La phrase nominale et la phrase verbale diffèrent de nature, et il est aussi vain de vouloir ramener l'une à l'autre, qu'il le serait de vouloir ramener le nom au verbe, et inversement.

Un même mot peut souvent servir de nom et de verbe dans les langues où le rôle du mot dans la phrase n'est pas indiqué par la forme du mot : l'anglais *love* a donné une idée du procédé, et a montré du même coup que cette identité de forme n'entraîne pas confusion de la catégorie du nom avec celle du verbe. Les mêmes éléments radicaux fournissent des noms ou des verbes suivant les affixes qu'on y adapte dans les langues où les mots ont des formations compliquées. Une même « racine » indo-européenne ou sémitique est un élément commun à la fois à des formes verbales et à des formes nominales ; en pareil cas, on parle souvent de racines verbales, parce que d'une racine telle que *ag-* qui forme le verbe latin *ago*, on tire aussi des noms tels que *actus* ou *actio* qui éveillent l'idée d'un procès, par association avec les formes verbales. Mais ces racines ne sont ni verbales ni nominales ; elles sont la partie du mot qui indique l'élément du sens commun à des verbes et à des noms. On voit là, de même que dans les mots susceptibles d'être à la fois noms et verbes, comment un même élément linguistique peut apparaître soit sous l'aspect nominal soit sous l'aspect verbal. Une « racine verbale » est, dans la structure morphologique des langues à flexion, ce qu'est dans les langues dénuées de toute flexion un mot capable d'être à la fois nom et verbe. Ceci n'emporte naturellement aucunes conséquences certaines pour la façon dont se sont constituées historiquement les racines indo-européennes ou sémitiques ; mais il est au moins vraisemblable que les « racines » sont des résidus de mots susceptibles d'être noms ou verbes suivant l'usage qui en est fait, comme anglais *love*, ou comme beaucoup de mots chinois.

*
* *

Les procès peuvent être envisagés de manières diverses ; il résulte de là que les formes verbales peuvent exprimer des caté-

gories diverses. Mais, comme, dans une langue donnée, le nombre des formes, et, par suite, celui des catégories est nécessairement très limité, le verbe présente les procès de manières différentes suivant les langues ; et les catégories exprimées par le verbe ne se laissent pas traduire exactement d'une langue à l'autre. Il est exceptionnel qu'on puisse traduire en français un verbe russe, ou en russe un verbe français sans rien laisser perdre de la valeur juste de chaque forme des deux langues. Même quand on réussit à faire une phrase russe qui équivaut exactement à une phrase française, le verbe russe exprime presque toujours une nuance qui n'est pas dans le verbe français, et inversement.

En donnant les mêmes noms à des formes grammaticales de langues diverses et en construisant autant que possible sur le même plan la grammaire de langues différentes, les grammairiens ont beaucoup péché ; ils ont répandu bien des idées fausses.

Tout au plus peut-on ramener les catégories qui apparaissent dans les diverses langues à un nombre restreint de types généraux. Mais on ne trouve pas partout les mêmes types ; des types qui, au fond, sont les mêmes, rendent, d'une langue à l'autre, des nuances assez différentes, de sorte que, même si l'on n'applique les mêmes noms qu'à des types semblables entre eux, néanmoins ces noms recouvrent encore des formes qui ont des valeurs sensiblement distinctes les unes des autres.

*
* *

L'habitude d'étudier la grammaire latine et l'influence qu'a exercée le modèle latin sur les auteurs de grammaires ont habitué à mettre au premier plan des catégories verbales celle du temps.

Rien de plus naturel au premier abord. Tout procès a lieu dans le temps, et il est naturel que la forme du verbe situe les procès dans le temps.

Toutefois, un détail frappe immédiatement : le temps comporte trois moments essentiels : le présent, le passé et l'avenir. Or, de ces trois moments, il arrive souvent que deux seulement soient exprimés par des formes grammaticales propres : le présent et le

passé. Le futur n'a souvent pas d'expression propre, ou, s'il en a une, c'est une expression compliquée et qui n'est pas parallèle à celles du présent et du passé. Ainsi l'allemand a un présent et un prétérit nets ; mais, au futur, le type *ich werde machen* est un tour lourd, gauche, qui s'emploie peu dans la langue parlée. L'anglais a des présents et des prétérits bien clairs, bien simples ; mais ses deux formes de futur sont d'un type autre, et elles expriment deux nuances de procès à venir, bien plutôt que le futur proprement dit : « je veux faire », ou « je dois faire ». Le germanique commun, dont les langues germaniques actuellement parlées sont des continuations, avait un présent et un prétérit, pas de futur. On n'a aucune raison de croire que l'indo-européen préhistorique, qui avait des moyens d'opposer le prétérit au présent, ait jamais possédé un vrai futur.

L'examen du développement linguistique confirme cette observation. Le futur est instable. Les langues romanes ont conservé le présent et le prétérit latins ; mais elles ont remplacé le futur latin par des formes nouvelles : *je chante*, *je chantais* continuent des formes latines ; *je chanterai* est une forme nouvelle, qui n'a rien à faire avec le latin *cantabo*, et qui s'est constituée dans le développement même du roman : elle signifiait d'abord « j'ai à chanter ». Aujourd'hui encore le français tend souvent à remplacer le futur, devenu tout abstrait, *je ferai*, par des formes expressives : *je vais faire*, *je veux faire*, *je dois faire*, etc.

L'opposition d'un présent et d'un prétérit est chose normale et durable ; l'opposition d'un futur au présent et au prétérit est exceptionnelle et sujette à changement.

Ce détail suffit à montrer que la catégorie « logique » du temps ne domine pas le verbe autant qu'on pourrait le croire au premier abord.

La façon dont le temps est envisagé varie d'ailleurs d'une langue à l'autre. Il y a des langues où plusieurs nuances de prétérit ont chacune leur marque spéciale, et il y en a d'autres où ne figure pas d'autre opposition que celle du présent et du prétérit, l'ancien germanique ou l'ancien slave, par exemple. Le français est le type d'une langue où des formes grammaticales distinctes

indiquent des moments finement nuancés du temps : il y a une unique forme de présent, puisque le présent, par nature, exclut toute nuance en tant qu'on ne considère que le temps proprement dit, mais il y a des prétérits multiples : l'un indique la simultanéité dans le passé : *je faisais cela autrefois, je faisais cela quand vous êtes arrivé* ; un autre indique simplement l'action passée : *j'ai fait cela* ; la langue parlée du centre de la France, tout autour de Paris, n'a plus que cette forme ; mais la langue écrite a conservé deux formes, *je fis* et *j'ai fait*, l'une exprimant purement et simplement que l'action de « faire » a eu lieu à un moment du passé, l'autre indiquant que l'action est réalisée : il y a ici un reste de distinction d'*aspect* dont on verra plus bas le principe ; une autre forme indique l'antériorité dans le passé : *j'avais fait cela auparavant, j'avais écrit cela quand vous êtes arrivé* ; aussi longtemps qu'a duré le prétérit simple, il y a eu une forme *j'eus fait* dont la nuance de prétérit antérieur concordait avec celle de *je fis* ; on tend aujourd'hui à dire : *j'ai eu fait*, par exemple *j'ai eu fait tout cela avant son arrivée*. Il y a également un futur et un futur antérieur. Il y a même un futur et un futur antérieur dans le passé : *je savais qu'il ferait cela, je savais qu'il aurait fait cela avant mon arrivée*. Hors des langues romanes, on ne trouverait guère pareille variété de nuances. Aussi, en réfléchissant sur le verbe, un Français doit-il se méfier de sa façon de concevoir le verbe comme une forme qui exprime avant tout le temps.

*
* *

Il y a une catégorie qui interfère souvent avec celle du « temps », c'est celle que l'on connaît en grammaire slave sous le nom d' « aspect », et dont le caractère répond aussi bien, sinon mieux, au caractère du verbe que la catégorie du « temps ».

La catégorie de l' « aspect », non moins variée que celle du « temps », embrasse tout ce qui est relatif à la durée et au degré d'achèvement des procès indiqués par les verbes.

En anglais, on distingue par exemple deux sortes de présents :

l'un, *I sing*, exprime simplement le fait que « je chante », l'autre, *I am singing*, signifie « je suis en train de chanter ». Et il y a deux prétérits correspondants : l'un, *I sang*, qui indique le fait que « j'ai chanté », dans le passé, et l'autre, *I was singing*, qui signifie « j'ai été, j'étais en train de chanter » dans le passé.

Le verbe slave est tout dominé par la notion de l'aspect. Un verbe slave n'est pas simple comme un verbe français ; il se compose toujours d'une paire de formes, l'une indiquant le procès, le fait pur et simple, et par suite souvent le procès achevé, l'autre indiquant le développement du procès. Celle des deux formes du verbe qui indique le procès purement et simplement est dite *perfective*, ainsi *pasti* « tomber » ; l'autre, qui indique le développement de l'action, est dite *imperfective : padati* « être en train de tomber ». C'est ainsi qu'en vieux slave *umrěti* « mourir » s'opposait à *umirati* « être en train de mourir ». L'importance et la constance de cette opposition de deux aspects caractérisent les langues slaves ; cette opposition s'est bien conservée dans toutes, surtout en russe, en tchèque, en polonais.

Mais il y a des langues où les oppositions sont plus complexes qu'elles ne sont en slave : le grec ancien distinguait non pas deux, mais trois « aspects », auxquels on donne des noms ou impropres ou signifiant peu de chose : l'aspect *duratif*, connu sous le nom très inexact de *présent*, qui répond à peu près à l'imperfectif slave, et qui fournit à la fois un présent proprement dit, soit *thnêskei* « il est en train de mourir », et un prétérit, dit imparfait, *ethnêske* « il était en train de mourir » — l'aspect *momentané*, connu sous le nom d'*aoriste*, qui répond en partie au perfectif slave, soit *ethane* « il est mort » ; l'aoriste ne comporte pas de temps présent, parce qu'il exprime un événement qui a eu lieu à un moment donné ; — l'aspect *achevé*, connu sous le nom de *parfait*, au moyen duquel on exprimait que le résultat du procès indiqué par le verbe est atteint ; le parfait grec ne répond qu'en partie au *perfectif* slave, soit *tethnêke* « il est mort » (non pas « il est mort à tel ou tel moment », mais « actuellement il est à l'état de mort ») ; le parfait s'emploie le plus souvent au présent, pour indiquer un résultat acquis au moment où parle le sujet ;

mais il en existe un prétérit, dit plus-que-parfait, soit *etethnêkei* « il était mort », pour exprimer un procès déjà réalisé dans le passé.

La catégorie de l'*aspect* est plus concrète que celle du *temps*, et, au cours de l'histoire des langues indo-européennes, on voit l'aspect perdre de l'importance, le temps en gagner.

Le latin offre à cet égard un fait caractéristique. Tandis que le grec conserve exactement, qu'il a peut-être même précisé et rendu plus rigoureuse l'opposition des trois aspects : duratif, momentané et parfait, le latin n'a gardé que deux aspects, un duratif, dit *infectum*, et un « perfectif », dit *perfectum*. Mais à chacun des deux aspects, il a donné à l'indicatif trois temps : un présent, un prétérit et un futur, si bien que l'on a le système suivant :

	Infectum.	Perfectum.
Présent. .	*dicit* : il est en train de dire,	*dixit* : il a dit.
Prétérit. .	*dicebat* : il était en train de dire,	*dixerat* : il avait dit.
Futur. . .	*dicet* : il sera en train de dire,	*dixerit* : il aura dit.

Le présent du perfectum *dixit* indique que l'action de « dire » est achevée au moment où l'on parle ; cette forme se prête également bien à raconter un fait passé : *dixi* « j'ai dit » (à un moment quelconque du passé) et à insister sur le fait qu'on a fini de parler : *dixi* « j'ai dit » (et je n'en dirai pas plus) ; certaines formes du présent du perfectum, comme *memini* « je me souviens » servent de présents. On a traduit le futur du perfectum *dixerit* par « il aura dit » ; mais la forme française rend mal la forme latine, et si l'on s'obstinait à rendre *dixerit* des textes latins par « il aura dit », on tomberait souvent dans de graves difficultés ; la forme latine signifie « il dira », étant entendu que l'action est envisagée comme achevée ; c'est par suite un futur plus fort que le futur *dicet*. (Pour ne pas compliquer un exposé, déjà trop touffu par nécessité, on laissera de côté ici la nuance fine qu'ajoute souvent à un verbe l'addition d'un « préverbe », c'est-à-dire l'opposition de *memini* et *commemini*, de *vinco* et *evinco*, etc.) — Les langues romanes n'ont pas gardé ce système, elles ont laissé

tomber tout ce qui avait valeur d'aspect, et elles n'ont gardé que la valeur temporelle. Le latin ancien se trouve donc représenter une transition de l'indo-européen, où dominait l'aspect, aux langues romanes actuelles, où domine le temps.— On observe ainsi trois moments successifs : l'indo-européen, où les divers radicaux d'un même verbe, celui du « présent », celui de l' « aoriste », et celui du « parfait », expriment des aspects, où la différence du présent et du passé n'est exprimée qu'accessoirement et d'une manière sommaire, parfois ambiguë, et où il n'y a pas de futur — le latin, où l'opposition d'une forme indiquant le procès en voie de développement, *l'infectum*, et d'une forme indiquant l'action achevée, le *perfectum*, se maintient, mais où il n'y a plus que deux aspects nets, et où à chacune des deux notions d'aspect se superpose l'expression nette des trois temps, présent, passé et futur — le français, où il ne reste rien de la notion d'aspect, et où le temps est rendu avec tout un luxe de nuances.

Les langues sémitiques, moins abstraites en général que les langues indo-européennes, opposent pour la plupart un « parfait » et un « imparfait », dont la valeur essentielle se rapporte à l'aspect ; et elles n'arrivent à indiquer le temps que par des procédés indirects, au moyen de formes qui se rapportent à la catégorie de l'aspect.

*
* *

Les besoins de l'expression amènent souvent à rétablir des oppositions d'aspect au fur et à mesure que s'efface par l'usage la valeur expressive des oppositions existantes.

L'histoire des formes du présent indo-européen est instructive à cet égard.

Le type dit du « présent » exprimait le procès considéré dans son développement, dans sa durée. Cette valeur est encore bien sensible dans les anciens textes indo-iraniens, et surtout en grec, et même en latin. Mais toutes les catégories qui rendent des notions expressives et concrètes plutôt que logiques tendent sans cesse à perdre de leur valeur, et la tendance à parler d'une

manière expressive et concrète en a sans cesse fait rétablir à la place de celles qui se sont effacées. C'est ainsi que le futur s'élimine, remplacé par des formes expressives qui signifient : « Je veux faire, je dois faire », etc. Les radicaux du « présent » indo-européen ont peu à peu perdu presque partout leur « valeur relative » ; ceci se reconnaît à ce qu'ils cessent de fournir un prétérit en même temps qu'un présent : un radical de « présent » indo-européen fournissait à la fois un « présent proprement dit » et un « imparfait », ainsi en grec *phérô* « je porte » et *épheron* « je portais », tandis qu'un présent slave, lituanien, germanique, irlandais n'est pas accompagné d'un imparfait, lié au radical du « présent ». C'est que le radical du « présent » indo-européen exprimait uniquement l'aspect, et que le radical du présent germanique, par exemple, exprime uniquement le temps.

Mais le besoin qu'on éprouve d'insister sur la durée d'une action qui se poursuit persiste. On y répond par divers procédés. Le persan se sert d'une particule *bē* pour indiquer la durée de l'action. Parmi les parlers arméniens, les uns ont tiré parti de l'usage de coupler des verbes qui existait dans la langue ; ils ont étendu l'emploi de *ka u* « il se tient et... », puis ont généralisé l'usage de ce groupe de mots devenu simple particule devant n'importe quel présent ; les autres ont recouru à une périphrase, et, au lieu de « il porte », ont dit « il est dans le porter » : *berum ē* ; actuellement, ces deux procédés sont fixés dans les divers parlers arméniens, et ils ont depuis longtemps perdu leur force ; mais on voit comment ils se sont créés et pour quelles fins expressives. Le slave a développé des imperfectifs qui expriment fortement la durée. L'anglais a fait sa forme composée : *I am reading* « je suis à lire ». Le français n'a encore aucun procédé grammatical fixé, et il doit recourir à des termes dont les éléments ne sont pas fondus et qui n'ont pas encore le caractère de formes grammaticales, comme : *je suis à lire*, *je suis en train de lire*, etc. ; (le berrichon a, en pareil cas, *je suis après labourer*, *al est après coudre*, « elle est en train de coudre », etc.) ; par suite, ces types ne sont pas aussi courants que le type anglais, *I am reading*, par exemple.

Les sujets parlants sentent aussi le besoin de marquer qu'un

procès est achevé. Au fur et à mesure qu'elles s'emploient, les formes grammaticales qui servent pour cette catégorie, les formes dites de « parfait », s'usent. Elles finissent par n'avoir d'autre valeur que celle d'un prétérit. Alors il faut créer des procédés nouveaux pour rendre le parfait. L'histoire des langues indo-européennes montre d'une manière frappante et l'usure du parfait indo-européen, et les créations successives auxquelles donne lieu cette usure, et le fait que des formes qui ont été faites pour rendre le parfait, c'est-à-dire l'*aspect* parfait, ont toujours de nouveau abouti à rendre le passé, c'est-à-dire le *temps* passé [1].

Ainsi, dès le latin classique, les formes de perfectum présent, telles que *dedi* « j'ai donné », *dixi* « j'ai dit », etc., tout en gardant leur nuance de sens propre, c'est-à-dire en continuant d'indiquer le procès achevé, servaient de temps historique, pour énoncer que tel ou tel procès a eu lieu dans le passé. Ces formes tendaient ainsi à perdre leur valeur de perfectum, et elles l'ont enfin si bien perdue que, dans toutes les langues romanes, elles n'ont fourni que des prétérits, et non plus des parfaits. Pour rendre le parfait, on a recouru à des procédés nouveaux. Le plus usuel consiste à prendre le verbe signifiant « avoir », et à y joindre un participe indiquant le procès achevé, soit *habeo illud factum* « je tiens cela fait, j'ai cela fait » ; il y a eu là d'abord deux mots indépendants l'un de l'autre, *habeo* et *factum* ; puis *habeo* et *factum* ont été rapprochés au point de vue du sens, et même au point de vue de la forme, et l'on est arrivé au type français : *j'ai fait*. A son tour, ce type a tendu à perdre sa valeur de parfait quand on a cessé de sentir *ai* et *fait* comme deux mots distincts ; aujourd'hui les trois éléments de *j'ai fait* ne sont pas plus distincts pour le sujet parlant que ne l'étaient en latin ceux de *feci* ; le fait que ces trois éléments sont encore susceptibles d'être séparés matériellement les uns des autres (ainsi dans : *je* ne vous l'*ai* jamais *fait*) n'a pas suffi à empêcher la fusion des valeurs. Aussi le type *j'ai fait* a-t-il passé au rôle de simple temps historique et a perdu toute valeur de parfait ; dès lors le

1. Voir ci-dessus, p. 141 et suiv.

type *je fis* était superflu, et le type nouvellement formé en roman *j'ai fait* en a pris la place en français parlé d'aujourd'hui, dans toute la France centrale autour de Paris.

Pareille histoire n'a rien que de courant. La façon dont se constitue un prétérit nouveau varie, dans le détail, d'une langue à l'autre ; mais, au fond, le procédé est partout le même.

Les langues germaniques, qui sont en contact avec les langues romanes, ont employé le même procédé : l'allemand dit *ich habe gemâcht*, tout comme le français dit *j'ai fait* ; de même que dans la France du centre, *j'ai fait* a pris toute la place de *je fis*, de même dans une très grande partie du domaine allemand, *ich habe gemacht* a éliminé *ich machte*. — On a retrouvé le même recours à un verbe « avoir » dans des parlers iraniens du centre de l'Asie tout à fait isolés et qui n'ont pu subir l'action du roman ou du germanique.

En slave, il a été procédé un peu autrement. Le nouveau parfait se compose d'un nom d'agent, accompagné au besoin du verbe « être ». « J'ai dit » s'exprime alors par *reklŭ jesmĭ* « je suis diseur ». Mais, dès avant l'époque historique, le type des noms en *-lŭ* indiquant l'agent ne s'employait pas ailleurs que dans cette expression verbale ; par suite *reklŭ jesmĭ* ne s'analysait plus en « je suis celui qui a dit » ; c'était un procédé pour indiquer que l'acte de dire était fait. Aussi, dès les plus anciens textes, la valeur de parfait de l'expression est-elle peu sensible. Dans une notable partie du groupe slave, et notamment en russe et en polonais, le parfait ancien a perdu ainsi sa valeur propre ; il est devenu un simple prétérit, et il a entièrement éliminé l'ancien prétérit simple : c'est, avec d'autres éléments linguistiques, le pendant exact des faits français et allemand.

L'arménien a procédé comme le slave. Mais il ne semble pas que la forme composée, faite pour exprimer le parfait, ait jusqu'à présent éliminé le prétérit simple dans la plupart des parlers.

L'iranien s'est servi d'un troisième procédé : un participe qui est intransitif ou passif suivant le verbe, et qui, comme en slave, fournit une phrase nominale ; ainsi, en vieux perse, *aita manâ*

krtam « cela a été fait par moi », *bûta a(h)mi* « j'ai été ». Par des arrangements successifs, ce type est devenu le seul type de prétérit du persan : *kardam* « j'ai fait », *bûdam* « j'ai été », etc.

Quel que soit le détail, il s'agit toujours d'un même type de développement. Le point de départ est le besoin d'exprimer le procès achevé, dans une langue où il y a déjà une forme pour exprimer simplement le prétérit. Le point d'arrivée est la constitution d'une nouvelle forme de prétérit qui souvent a remplacé tout à fait l'ancienne.

Ce développement, qui se répète sans cesse, montre comment la catégorie expressive et concrète de l'aspect fournit, en perdant par l'usage son caractère expressif et concret, un moyen de rendre la catégorie abstraite du temps.

*
* *

Sous le nom de *modes* on entend les formes au moyen desquelles est indiquée l'attitude mentale du sujet parlant par rapport au procès indiqué par le verbe.

On pourrait s'attendre à ce que la différence entre l'énonciation des faits et le commandement détermine deux catégories grammaticales tranchées. Or, il se trouve que souvent les différences entre les formes qui servent à énoncer et celles qui servent à commander sont ou nulles ou minimes. Ainsi, en grec, c'est une même forme *phérete* qui sert à indiquer le fait : « vous portez », et le commandement : « portez ». En français, c'est exactement la même forme verbale qui sert à énoncer et à ordonner ; seulement dans un cas il y a pronom préfixé, et dans l'autre non : *chante* ou *tu chantes* (l'*s* est purement graphique), *viens* et *tu viens*, etc. En latin, il y a une légère différence entre les formes qui servent à énoncer et celles qui servent à commander, mais cette différence ne porte que sur la désinence, nullement sur le radical : *canta* « chante », et *cantate* « chantez » ne se distinguent que légèrement de *cantas* « tu chantes », *cantatis* « vous chantez ». L'absence ou la faiblesse de la distinction entre les formes qui servent à deux usages aussi distincts que

l'énonciation et le commandement vient de ce que, dans le discours parlé, la situation des interlocuteurs et le ton du discours indiquent assez si l'on énonce ou si l'on commande.

On remarquera seulement, en passant, que la forme qui sert à ordonner, à la 2e personne du singulier, ou bien se confond avec le radical du verbe, ou s'en distingue assez peu : ainsi en français *chante*, *viens*, *cours*, etc. C'est que l'on parle surtout pour obtenir quelque action de l'interlocuteur, et que, au fond, l'impératif est la forme essentielle du verbe. Les textes écrits, sur lesquels les linguistes opèrent le plus souvent, donnent à cet égard une idée fausse de la réalité.

En revanche, — et surtout chez des demi-civilisés ou chez des hommes de faible culture, — le sujet parlant tend à marquer son attitude vis-à-vis du procès dont il s'agit. Il tient à indiquer expressément le fait ou le désir, la volonté, la certitude ou la possibilité. Ici encore le développement des formes au cours de l'histoire des langues indo-européennes est instructif.

L'indo-européen commun, organe d'un peuple simplement demi-civilisé, distinguait un *indicatif*, au moyen duquel on énonçait le procès, de deux « modes » dont la forme apparaît comme dérivée par rapport à celle de l'indicatif : le *subjonctif*, qui servait à exprimer une volonté, et l'*optatif*, qui servait à indiquer une possibilité ou un désir. Mais cette distinction de trois modes ne s'observe plus que dans les deux langues de la famille indo-européenne attestées à la date la plus ancienne et sous la forme la plus archaïque : le grec et l'indo-iranien. Car, dès le début, la langue tend à ne plus opposer que deux modes : l'un, l'indicatif, énonçant le procès comme un fait, l'autre, dit subjonctif, de forme dérivée, énonçant le procès avec quelque considération subjective : volonté, désir, possibilité. En grec et en iranien, l'optatif s'est éliminé d'assez bonne heure (l'optatif grec était sorti de l'usage de la langue parlée dès le 1er siècle après J.-C.) ; en sanskrit, c'est le subjonctif qui a succombé, bien avant le début de l'ère chrétienne à ce qu'il semble. Et partout, il n'est resté qu'un seul mode à valeur subjective. On n'est donc pas surpris de constater que des langues comme le latin, l'irlan-

dais, les dialectes germaniques offrent un seul mode distinct de l'indicatif.

Du reste, avec le temps et le progrès de la culture, le mode unique opposé à l'indicatif change de rôle. En indo-européen, chaque mode avait sa valeur sémantique propre. Peu à peu, au fur et à mesure que la structure des phrases se complique et qu'à l'usage dominant des phrases simplement coordonnées et des phrases relatives, se superpose un système compliqué de phrases subordonnées, le mode à forme dérivée sert surtout à caractériser certains types de subordonnées et certaines nuances de sens chez les subordonnées. Déjà en latin, le subjonctif est avant tout un outil de la subordination ; il sert à caractériser certains types de subordonnées ou certaines valeurs spéciales chez les subordonnées : *hominem quaero qui ueniat* « je cherche un homme pour qu'il vienne » diffère absolument de *hominem quaero qui uenit* « je cherche l'homme qui vient ». On retrouve facilement dans les phrases où figure le subjonctif latin, la nuance de la volonté, celle de la possibilité, mais tout cela lié la plupart du temps à certains types de phrases dépendantes. Il est relativement rare que le subjonctif latin s'emploie dans une phrase principale, avec sa valeur propre ; la grande majorité des exemples figure dans des phrases dépendantes.

Le français, qui oppose *je crois qu'il le veut* à *je ne crois pas qu'il le veuille*, par exemple, et *je sais qu'il vient* à *je veux qu'il vienne*, montre bien comment le subjonctif, qui d'abord exprimait la sensibilité du sujet parlant, est aujourd'hui le moyen de caractériser certains types de phrases. Aussi le français a-t-il constitué un nouveau type modal, qui a une valeur surtout grammaticale, le conditionnel : *je ferais volontiers ceci si vous vouliez*. Parfois le conditionnel sert à indiquer la possibilité, dans des conditions assez pareilles à celles où l'on rencontre certains optatifs grecs anciens : *je ferais volontiers ceci* ; mais son rôle le plus ordinaire est de figurer dans des phrases exprimant une condition.

On remarquera ce qui se passe dans certaines langues indo-européennes qui ont derrière elles un long développement, puisque les plus anciens monuments sont du IX^e^ siècle après

Jésus-Christ pour le slave, du XVI^e^ pour le groupe baltique, et où par suite le mode a pu s'altérer beaucoup ; elles ont perdu jusqu'à l'opposition d'un « subjonctif » et d'un indicatif, sans doute parce que, demeurant à un stade archaïque et servant à des populations relativement arriérées, elles sont allées jusqu'au bout de l'évolution avant d'avoir beaucoup développé la subordination.

Ainsi le rôle du « mode » a varié, au fur et à mesure que le sujet parlant a moins éprouvé le besoin de marquer, dans les formes verbales qu'il emploie, son attitude mentale vis-à-vis des procès indiqués.

*
* *

Il y a tout un autre groupe de catégories qui a une grande importance pour la théorie du verbe ; ce sont celles qui se rapportent au rôle, par rapport au verbe, de la notion à laquelle s'applique le procès indiqué, de ce que l'on appelle le « sujet ». On est amené à distinguer des verbes *actifs* et *passifs*, *transitifs* et *intransitifs*, *moyens* ou *réfléchis*, *causatifs*, *désidératifs*, *intensifs*. Comme il s'agit de valeurs sémantiques de caractère ou concret ou sentimental, la plupart de ces oppositions ont tendu, avec le temps, à s'atténuer ou à disparaître.

On nomme *sujet* la notion à laquelle s'applique le procès indiqué par le verbe : *Pierre vient*, *Pierre dort*, *les blés mûrissent*, *le monument s'élève là*, etc. Les formes fléchies à marque personnelle comprennent elles-mêmes leur sujet : latin *canto*, *cantas*, *cantamus*, *cantatis* ont leur sujet dans la forme même du verbe ; et, comme, en français, les anciens pronoms *je* ou *tu* sont devenus des éléments de flexion, on peut dire que de même *je chante*, *tu chantes* comprennent leur sujet. S'il s'agit d'une personne qui est présente à l'esprit des interlocuteurs, la forme *cantat* se suffit en latin, et, en français, *il chante*, *elle chante*, où *il*, *elle* ne sont que des outils grammaticaux, et non plus des pronoms autonomes ; ceci est si vrai que, comme on l'a indiqué ci-dessus, ces anciens pronoms tendent à faire corps avec les verbes : pour annoncer que de l'eau qui a été mise à chauffer est bouillante, un Berr-

chon dit couramment : « *ton eau, a bout* », c'est-à-dire « ton eau, elle bout », là où le français littéraire aurait *ton eau bout*.

Le procès indiqué par le verbe peut se présenter de façons très diverses par rapport au « sujet » ; on appelle *voix* la catégorie qui se rapporte à ce type de distinctions.

Le procès peut être présenté comme un acte du sujet : *Pierre vient*, *Pierre mange* ; c'est ce qu'exprime un verbe *actif*, en prenant le mot « actif » au sens large. L'acte peut être exprimé complètement par le verbe, ou bien il peut porter sur quelque objet qui doit être exprimé séparément : *Pierre frappe Paul*, *Pierre mange du pain*. Les verbes qui comportent un objet, sont dits *transitifs*, et ceux qui n'en comportent pas, *intransitifs*. Certains verbes sont ordinairement intransitifs ; ainsi *dormir*, *venir*, etc. ; d'autres ne le sont qu'en certains cas, ainsi *manger* ; et, dans ce dernier cas, on distingue un emploi transitif et un emploi intransitif. Dans les anciennes langues indo-européennes, les verbes qui sont susceptibles de l'emploi transitif, l'étaient en même temps d'un emploi intransitif, dit *absolu* : un verbe qui signifie « je laisse » pouvait aussi s'employer intransitivement, et alors le sens était « je reste », l'action de « laisser » s'appliquant au sujet lui-même ; un verbe qui signifiait « je tiens » pouvait, sans complément, signifier « je me tiens », l'action de tenir s'appliquant au sujet ; ainsi, en latin, *uerte* signifie « tourne » (quelque chose) et « tourne-toi ». L'emploi absolu est particulièrement courant là où existe une détermination qui indique comment se fait l'action indiquée par le verbe ; en français on ne peut guère dire *il frappe*, sans complément, à moins qu'on ne sache ce qui est frappé ; mais on peut dire *il frappe fort* ; en grec où *ekhô* signifie « je tiens, j'ai (quelque chose) », on dit par exemple *kakôs ekhô* « je me tiens mal, je suis en mauvais état ». Le procédé qui consiste à employer les mêmes verbes transitivement et intransitivement peut avoir plus ou moins d'extension suivant les langues, mais la nature des choses fait qu'il existe à peu près nécessairement.

En revanche, l'indo-européen avait une particularité singulière, si singulière qu'elle s'est éliminée presque partout, et que, si l'on

n'avait pas le grec ancien et l'ancien indo-iranien, on ne saurait s'en faire une idée juste. C'est la distinction de *l'actif* et du *moyen*. *L'actif* servait à indiquer un procès auquel le « sujet » n'est pas particulièrement intéressé, le *moyen* un procès auquel il avait un intérêt propre de quelque manière : ainsi en sanskrit védique, du prêtre qui fait un sacrifice pour quelqu'un qui le lui fait faire, on dit à l'actif *yajati* ; mais, du personnage qui a commandé le sacrifice au prêtre et qui est lui-même un des agents du sacrifice, on dit au moyen *yajate* ; l'opposition de l'actif *thyei* « il sacrifie » et du moyen *thyetai* « il sacrifie pour lui-même » a la même valeur en grec ancien. Il va de soi que cette opposition n'a pas lieu d'exister pour tous les verbes. Tel verbe signifiant « aller » par exemple, ne s'emploie qu'à l'actif, ainsi en sanskrit *emi* « je vais » et en grec *eimi* ; tel autre signifiant « avoir une agitation mentale, penser », indiquant par conséquent une activité intérieure du sujet, est toujours moyen, ainsi *manye* « je pense » en sanskrit, *mainomai* « j'ai une agitation mentale » en grec. Il est presque toujours facile de rendre compte de l'emploi des désinences actives ou moyennes en ancien indo-iranien et en ancien grec. Mais une opposition de ce genre avait un caractère trop spécial et trop concret, elle entraînait d'ailleurs une trop grande variété de formes pour durer : un ancien verbe grec ou indo-iranien a au complet une double série de désinences, l'une pour l'actif, l'autre pour le moyen, et cela suffit à doubler le nombre, déjà très grand sans cela, des formes de la conjugaison dans ces langues. Aussi la distinction s'est-elle éliminée de bonne heure ; il y a des langues où il en subsiste des traces, d'autres où elle est tout à fait disparue ; en grec même et en indo-iranien, elle a cessé d'exister au cours de l'époque historique.

La possibilité d'employer les verbes actifs absolument et le fait que les désinences moyennes soulignent au besoin cette valeur absolue ont dispensé l'indo-européen commun d'avoir un passif. On croit souvent que le passif est une forme du verbe où le sujet du verbe est indiqué comme subissant une action exercée par un agent : *Paul est battu par Pierre* ; les expressions de ce genre

se rencontrent en effet ; mais ce sont des tours souvent artificiels, et en tout cas relativement rares. Le vrai rôle du passif est d'exprimer le procès là où l'agent n'est pas considéré. Le latin a *dicit* « il dit », quand on pense à quelqu'un qui parle, et, à côté, *dicitur* « il est dit », pour signifier « on dit » ; *dicitur* équivaut à *dicunt* « ils disent », avec sujet indéfini, équivalant par suite aussi à « on dit ». Pour le sémitique, où le passif est une partie constitutive du système verbal, cette valeur a été reconnue par les grammairiens arabes. Si le passif n'était qu'un renversement de l'expression active, il serait au fond superflu. Ce qui donne au passif son utilité, c'est que, au lieu de présenter le procès comme résultant de l'intervention d'un agent, il le présente en lui-même, sans aucune notion étrangère. Si, près d'un passif, on marque l'agent, c'est comme un point de départ de l'action, non comme un agent proprement dit : le latin *occiditur a Marco* « il est tué par Marcus » signifie proprement : « il est tué » et « le point de départ de ce fait est Marcus » ; les expressions analogues du slave et du grec le montrent tout aussi clairement. Le passif peut n'exister que sous forme impersonnelle : le passif *dicitur* « il est dit », au sens de « on dit », est courant en latin. Le passif existe du reste surtout à la 3e personne.

Si l'indo-européen avait une opposition de l'*actif* et du *moyen*, et s'il n'avait pas d'opposition de l'*actif* et du *passif*, c'est que c'est une langue où les procès sont présentés en général d'une manière active, et comme résultant de l'intervention d'un agent plus ou moins personnel, plus ou moins défini. Quand un Français d'aujourd'hui dit « il vente », on parle à juste titre d'un verbe « impersonnel » : la forme employée signifie simplement que « le vent souffle », et on ne fait allusion à l'action d'aucune personnalité définie ; mais quand un poète de l'époque védique disait *vâti*, il voulait dire que *vâyu*, qui est le vent, mais qui est aussi un agent ayant une personnalité, un « dieu », exerce son activité spécifique de « venter ». Linguistiquement, les deux constructions ont l'air toutes pareilles ; mais elles expriment deux mentalités absolument différentes l'une de l'autre. — Or, dans une langue où tout procès est présenté normalement

comme résultant de l'action d'un être plus ou moins personnel, il est naturel de distinguer si le procès a une relation spéciale avec celui qui le produit ; mais une catégorie grammaticale spéciale pour indiquer le procès sans considération d'agent est assez superflue. En somme, l'opposition de l'actif et du moyen caractérise l'indo-européen.

Le moyen pouvait indiquer que l'action du sujet a pour objet le sujet même : la forme moyenne *louetai* du grec signifie « il se lave », par opposition à *louei* qui signifie « il lave ». Mais, pas plus qu'il n'avait un vrai passif, l'indo-européen n'avait un vrai réfléchi. On a été en général amené plus ou moins vite à indiquer le verbe réfléchi à l'aide d'un mot spécial, tel que le pronom réfléchi. Et, cette forme à pronom réfléchi, perdant progressivement de sa valeur par l'emploi, a fini par équivaloir au moyen ancien et par le remplacer ; c'est ce qu'on observe notamment en slave. En français, surtout au moyen âge, le verbe réfléchi a ainsi servi à noter une action en tant qu'elle intéresse particulièrement le sujet : de même que le latin disait *moritur*, avec sa forme « déponente » qui continue l'ancienne forme moyenne avec un mélange d'autres formes, le français a fait *il se meurt*, à côté de *il meurt* ; la forme est plus expressive que le simple *il meurt* et montre mieux la réalisation du procès, parce que la relation du procès avec le sujet est mise en évidence.

L'agent d'une action peut ne pas la faire par lui-même ; il la réalise souvent par l'intermédiaire d'un autre ; il faut donc que le verbe comporte un factitif. Le factitif peut être obtenu par un mot qui a perdu son sens propre et qui est devenu, par là, un simple outil grammatical, comme « faire » dans le français *il fait venir* et, ce qui est bien caractéristique, *il fait faire*, ou comme *lassen* en allemand. Mais très souvent aussi, il y a des fonctions spéciales au factitif : à côté de *bharati* « il porte », le sanskrit a *bharayati* « il fait porter » ; en face de *faran* « aller (en véhicule) », le vieux haut allemand a *fuoran* « faire aller en véhicule, conduire » (allemand moderne *führen*), et ainsi souvent.

On peut aussi indiquer qu'on a l'intention de réaliser un pro-

cès : une langue telle que l'indo-européen où, on l'a vu, les procès sont présentés « activement », a été amenée à constituer un *désidératif* : le latin, par exemple, a *uiso* « je désire voir », à côté de *uideo* « je vois », et beaucoup de formes analogues, qui continuaient un usage indo-européen très considérable.

Enfin, bien que ce type de formation se rapporte moins directement au rôle du sujet par rapport au verbe, il faut rappeler ici qu'il y a lieu de marquer si l'action se répète, si elle est faite avec intensité. Il existe des formations *itératives*, *intensives*. On peut marquer l'intensité par un redoublement plus ou moins complet de l'élément radical ; ainsi, en grec, un verbe tel que *marmairô* « je brille » est intensif. Un autre procédé, qui tient une grande place dans les langues sémitiques, mais qui se rencontre aussi dans les langues indo-européennes, et surtout, semble-t-il, dans la manière familière et populaire d'employer les anciennes langues indo-européennes, consiste à géminer une consonne intérieure : à côté de *glutire*, le latin a *gluttire* « engloutir » pour noter le procès avec intensité ; le français *en-gloutir*, où la gémination n'est pas conservée, conserve encore la force de sens que cette gémination a donnée.

Dans les langues des peuples incomplètement civilisés où les catégories grammaticales se rapportent à des notions en partie concrètes et où elles ont souvent une valeur expressive et sentimentale, les formes factitives, désidératives, intensives sont naturelles ; elles tendent à s'éliminer au fur et à mesure que, avec le progrès de la civilisation, les catégories prennent un caractère plus abstrait. Elles tendent à se refaire plus ou moins dans les parlers populaires.

*
* *

En somme, le verbe exprimant un procès, les catégories principales que les langues ont été amenées à créer sont celles de la *personne* (comprenant indirectement celle du *nombre*), celles du *temps* et de l'*aspect*, celle du *mode* et celle de la *voix*. Le progrès de la civilisation met en évidence le *temps* ; il tend à éliminer les catégories à valeur concrète ou expressive, et à donner aux catégories abstraites une importance de plus en plus grande.

LE
GENRE GRAMMATICAL
ET
L'ÉLIMINATION DE LA FLEXION [1]

On suit le développement de l'indo-européen depuis une époque préhistorique, où il n'y avait qu'un seul idiome indo-européen, jusqu'à l'époque actuelle, où cet idiome s'est différencié en un grand nombre de langues distinctes très peu semblables entre elles. Ce développement, dont l'étendue n'est pas inférieure à quatre millénaires et dont la variété est extrême, fournit l'expérience linguistique la plus complète qu'on possède. Il a été étudié en détail, souvent avec minutie. Mais on n'a pas réussi encore à en tirer des conclusions générales aussi précises et aussi riches que le ferait attendre la masse des faits patiemment assemblés, discutés, critiqués et classés.

Les conditions dont dépendent les changements de formes grammaticales sont si nombreuses et si complexes qu'il est le plus souvent impossible de déterminer d'où procède un changement donné. Mais le rapprochement des faits permet de dégager certaines tendances.

Un trait est commun à toutes les langues indo-européennes : toutes ont simplifié la flexion indo-européenne. Même celles qui sont connues à la date la plus ancienne, comme le sanskrit védique, le dialecte iranien des gâthâs de l'Avesta, la langue homérique, ou celles qui, à des époques postérieures, ont conservé un aspect archaïque, comme le vieux slave des premiers traducteurs ou le lituanien du XVI[e] siècle et même le lituanien

1. *Scientia* (*Rivista di scienza*), vol. XXV (1919), n° LXXXVI, 6.

moderne, toutes, elles ont atténué la luxuriance de la morphologie indo-européenne. Cette tendance n'a pas cessé d'agir à l'époque historique : car partout, entre les plus anciens états de langue attestés et les états postérieurs, les restes subsistants de l'ancienne morphologie indo-européenne se sont réduits progressivement.

Comme l'étendue et la nature de la réduction varient beaucoup d'une langue indo-européenne à l'autre, on peut faire un départ entre ce qui provient de conditions communes à tout le groupe et ce qui provient de conditions spéciales à chaque langue. Les changements qui se produisent partout ou presque partout sont ceux qui sont dus à des conditions générales.

Les conditions des changements sont, les unes universelles et constantes, les autres liées à des événements historiques. Les unes et les autres ont conduit à simplifier la flexion indo-européenne.

Les formes grammaticales vont sans cesse se normalisant : il est naturel d'exprimer constamment de la même manière une même catégorie grammaticale dans une même langue. Par suite, si des accidents divers, souvent de nature phonétique, n'intervenaient pas, les langues arriveraient avec le temps à une régularité parfaite, c'est-à-dire à l'unité d'expression de chaque catégorie grammaticale. Certaines langues sont même assez près de cet idéal, ainsi le finnois. Or, l'indo-européen commun exprimait les mêmes catégories de façons très diverses. Par exemple, les formes casuelles ont des caractéristiques distinctes pour chacun des trois nombres exprimés par la langue : singulier, pluriel et duel, et aussi en partie suivant qu'il s'agit d'un nom neutre ou d'un nom masculin ou féminin. De plus, les caractéristiques variaient suivant la forme du nom. Il résulte de là que la plupart des formes casuelles ont huit ou dix caractéristiques distinctes. Cette extrême complication qui est le caractère le plus frappant de l'indo-européen commun, et dont les restes dénoncent dès l'abord une langue indo-européenne au linguiste exercé, va contre le principe même de la morphologie.

Toutefois l'expérience montre qu'une morphologie même très

compliquée se maintient avec une grande obstination pourvu que le milieu linguistique soit stable et homogène. Mais les langues indo-européennes se sont sans cesse étendues sur des domaines nouveaux ; sans cesse, par suite, elles ont été acquises par des populations qui antérieurement employaient d'autres langues. Or, il va de soi qu'une population qui acquiert une langue nouvelle a peine à s'assimiler une morphologie complexe et délicate. Et, d'autre part, durant tout le temps où sont juxtaposées deux populations de langue différente sur un même domaine, il y a des individus qui parlent d'une manière imparfaite l'une des deux langues et qui introduisent le trouble dans les parties délicates ou peu stables de la morphologie. Toute extension de langue a donc pour effet de simplifier le système et d'y supprimer des distinctions plus ou moins superflues.

On s'explique par là le fait que, et dans le groupe indo-européen et dans le groupe sémitique, les complications de la flexion aient tendu à s'éliminer au fur et à mesure que les langues du groupe s'étendaient à des populations nouvelles.

Un fait matériel a contribué à la réduction de la flexion : les caractéristiques de la flexion se trouvaient surtout dans la fin du mot indo-européen ; or, cette fin est en général prononcée faiblement et tend à s'abréger ou même à disparaître. Les caractéristiques s'altèrent donc, ou s'effacent, au cours du développement des langues indo-européennes. Cette circonstance, purement phonétique, indépendante du sens à exprimer, a agi dans le même sens que la tendance à normaliser inhérente à toute morphologie et que la difficulté éprouvée par des populations nouvelles à s'assimiler une grammaire compliquée.

*
* *

Le sens dans lequel se fait la simplification est déterminé par la mentalité des sujets parlants. Le progrès de la civilisation détermine un progrès de la pensée abstraite, et, au cours du développement des langues indo-européennes, on voit les catégories grammaticales concrètes disparaître peu à peu, tandis que

les catégories qui répondent bien aux catégories abstraites de la pensée se maintiennent ou se développent. Comme le progrès n'a pas été partout également rapide, les effets de cette tendance peuvent se laisser isoler des changements dus à d'autres faits.

Or, le genre grammatical est l'une des catégories grammaticales les moins logiques et les plus inattendues.

On peut distinguer les êtres animés et les choses. C'est cette distinction d'un genre « animé » et d'un genre « inanimé » qu'exprime l'opposition du masculin-féminin, d'une part, du neutre, de l'autre, dans les langues indo-européennes. Mais elle ne s'applique que dans un nombre de cas restreint, et elle n'a pas été poursuivie d'une manière complète et rigoureuse. Du reste, les langues indo-européennes ne l'expriment que pour le sujet de la phrase et pour le complément direct (nominatif et accusatif). Dans tout le reste de la flexion, la distinction n'apparaît pas : le neutre ne se distingue du masculin-féminin qu'à ces deux cas.

On peut, d'autre part, distinguer les mâles et les femelles : là où la distinction du masculin et du féminin a une valeur, c'est la réalité qu'elle exprime. Le démonstratif sanskrit *sa* signifiait « cet homme », par opposition à *sâ* qui désignait « cette femme ». Cette opposition n'a de sens que pour les hommes et pour quelques animaux où la distinction des sexes a une importance. Pour les noms qui désignent toutes les autres notions, elle n'a aucune signification. Or, les adjectifs indo-européens affectent régulièrement deux formes, ainsi en sanskrit *nava-* et *navâ-*, et l'on emploie régulièrement l'une de ces deux formes suivant le substantif auquel se rapporte l'adjectif sans que le sexe intervienne directement en rien. Les substantifs avec lesquels on emploie la flexion du type *nava-* sont dits masculins ou neutres, suivant la manière dont ils se fléchissent au nominatif et à l'accusatif ; ceux avec lesquels on emploie la flexion du type *navâ-* sont dits féminins.

Cette distinction qui traverse toute la langue ne répond plus à rien dans la grande majorité des cas : par exemple, certains noms abstraits sont masculins, d'autres féminins et d'autres

neutres, sans qu'on voie la raison de ces différences. Le nom de certains objets est masculin, celui de certains autres féminin, et celui de certains autres neutre sans raison visible. En français où la distinction du masculin et du féminin subsiste encore, on ne saurait dire pourquoi le *siège* est masculin et la *chaise* féminine, pourquoi le *soleil* est masculin et la *lune* féminine. En allemand, le nom *Sonne* du « soleil » est féminin et le nom *Mond* de la « lune » masculin; en russe, *solnce* « soleil » est neutre, et des deux noms de la « lune », l'un, *mêsiac*, est masculin, et l'autre, *luna*, féminin, et tout cela sans raison visible. Derrière tous ces faits, on entrevoit l'action de conceptions anciennes, dont l'article suivant donnera un aperçu; mais, de très bonne heure, ces conceptions ont cessé d'être dominantes.

Presque toujours dénuée de sens, au moins à l'époque historique, la distinction des genres grammaticaux était de plus incohérente dans la forme dès l'époque indo-européenne.

La distinction du masculin-féminin, d'une part, du neutre, de l'autre, n'existait, on l'a déjà noté, qu'à deux cas, au nominatif et à l'accusatif. Mais à ces deux cas, elle s'exprimait par la flexion même : à un nominatif-accusatif singulier neutre comme *dulce* « doux » du latin s'opposent un nominatif *dulcis* et un accusatif *dulcem* qui sont masculins-féminins. Au contraire, l'opposition du masculin et du féminin s'exprimait, non par une différence de la flexion, mais par une différence du mot lui-même; en latin la différence entre l'accusatif masculin *bonum* et l'accusatif féminin *bonam* n'est pas dans la désinence, qui est *-m* dans les deux cas; elle est dans le fait que, au masculin, le mot fléchi est *bono-* (*bonum* représente un ancien *dwenom*) et, au féminin, le mot fléchi est *bona-* (ancien *dwenâ-*). Dans une langue comme le latin, cette différence entre le mode d'expression du masculin et du féminin et celui du masculin-féminin et du neutre est souvent masquée ; mais elle était nettement tranchée en indo-européen.

De plus, la distinction du masculin et du féminin ne s'exprimait pas dans le substantif, mais seulement dans les adjectifs qui s'y rapportaient. Un substantif en *-â-* pouvait être ou masculin ou féminin; et le latin par exemple continue à présenter des mas-

culins comme *scriba* à côté de féminins comme *lupa* « louve » ou comme *toga* « toge ». Un mot grec comme *hippos* désigne à la fois le « cheval » et la « jument », et les distinctions comme celle de *lupus* « loup » et de *lupa* « louve » en latin sont sûrement secondaires et récentes. Les procédés par lesquels on distingue les substantifs désignant des mâles de ceux désignant des femelles se sont constitués seulement au cours de l'histoire propre de chacune des langues indo-européennes et ils sont toujours restés incomplets.

En principe, on reconnaît un masculin comme *scriba* uniquement à ce que l'adjectif qui s'y rapporte est de la forme *bonus, bonum*, tandis que l'adjectif qui se rapporte à des féminins comme *lupa* ou *toga* est de la forme *bona, bonam*. Au contraire un nominatif neutre de la forme *iugum* « joug » se distingue nettement par lui-même d'un nominatif masculin de la forme *oculus* « œil » ou féminin de la forme *fagus* « hêtre ». A ce point de vue encore, il y a discordance de forme dans l'expression du genre.

Dans la mesure où, pour un même mot, une forme désignant la femelle s'oppose à une forme désignant le mâle, le nom de la femelle apparaît comme un dérivé de celui du mâle ; par exemple, en regard du nom qui désigne le mâle qui engendre, sanskrit *janitar-*, *janitr-*, latin *genitor-*, la forme qui désigne la femelle qui engendre est sanskrit *janitr-ī-*, latin *genetr-īc-*, avec des suffixes de dérivation.

La catégorie du genre était donc de celles qui devaient tendre à se réduire ou à s'éliminer au cours du développement de la langue. Elle avait contre elle à la fois de s'exprimer d'une manière incohérente, de n'être pas cohérente pour le sens, d'avoir un sens purement concret là où elle en avait un, et, le plus souvent, de n'en avoir aucun. Néanmoins elle a subsisté au complet jusqu'à présent sur une grande partie du domaine indo-européen et, dans d'autres langues, pourtant très avancées à d'autres égards, elle subsiste en partie.

Au contraire, la distinction d'un nombre duel distinct du nombre pluriel s'est éliminée partout avec le progrès de la civilisation.

Si les plus anciens textes indo-iraniens offrent encore le duel, on voit que la distinction avait disparu de l'usage courant dès avant l'époque chrétienne dans les parlers indo-iraniens sur lesquels on a des données. En Grèce, le duel a disparu, dès avant les plus anciens textes, dans les régions coloniales où le développement est relativement rapide, dès le IVe siècle dans des régions continentales comme l'Attique, où l'avancement a été plus tardif. Chez les Slaves, atteints relativement tard par la civilisation, le duel était encore en plein usage au IXe siècle après J.-C., et l'emploi n'en a pas encore disparu tout à fait chez les Slovènes par exemple. C'est que le duel avait dès l'indo-européen une flexion moins richement différenciée que le singulier ou le pluriel, que l'emploi en était naturellement plus restreint et que ce nombre n'avait pas de rôle particulier dans la structure de la phrase ; par suite il offrait peu de résistance à la tendance qui provoque la suppression des catégories grammaticales de caractère concret.

L'élimination de la flexion casuelle dont on observe les progrès au cours de l'histoire de presque toutes les langues indo-européennes va de pair avec les transformations du système de la langue. En indo-européen commun, et encore en ancien indo-iranien, en grec ancien, en latin ancien, en slave ancien (et jusque en russe ou en polonais actuels par exemple), chaque mot indiquait par les formes variées qu'il affectait le rôle qu'il jouait dans la phrase, et il se suffisait à lui-même. Dans la plupart des formes modernes des langues indo-européennes, au contraire, le mot a une forme fixée, et le rôle qu'il joue est indiqué soit par sa place vis-à-vis des autres mots de la phrase, soit par des mots accessoires : il ne se suffit pas à lui-même. Seules ont échappé à cette tendance les langues de caractère archaïque, le lituanien et la plupart des dialectes slaves, ou une langue qui s'est développée en un sens particulier, l'arménien.

L'élimination du genre ne tient pas à des tendances aussi générales, et par suite elle n'a pas été aussi complète.

*
* *

Néanmoins celles des langues dont le développement est poussé très avant ont ou réduit ou supprimé le genre grammatical.

Tandis que l'allemand, qui ne s'est pas encore débarrassé entièrement de la flexion casuelle et qui continue à distinguer inutilement un nominatif, un accusatif, un génitif et un datif, distingue aussi le masculin, le féminin et le neutre, et que les langues scandinaves ont également gardé au moins la distinction du masculin-féminin et du neutre, l'anglais, plus avancé, a profité de la destruction totale de la fin de mot où figurait la marque du genre pour écarter une distinction inutile. Le seul débris de genre qu'il ait conservé consiste dans l'emploi du pronom *he* « il » en parlant d'hommes, *she* « elle » en parlant de femmes, *it* « ceci » en parlant d'animaux et de choses ; cette différence constante, propre au pronom, n'a plus aucun des anciens inconvénients du genre, et l'anglais a réalisé sur ce point un progrès décisif.

Les langues romanes, qui toutes ont conservé de quelque manière la distinction de l'*o* et de l'*a* en fin de mot, soit expressément comme l'italien qui oppose *buono* à *buona*, soit indirectement comme le français qui oppose *bon* à *bonne*, ont gardé aussi la distinction du masculin et du féminin, tandis que le neutre se perdait dans le masculin. L'une des distinctions de genre était ainsi maintenue. Mais l'élimination du neutre montre la tendance à éliminer une catégorie dénuée de sens, et qui sert seulement, par les faits d'accord, à la structure de la phrase.

Une opposition comme celle de *un beau siège* et *une belle chaise* avait un appui trop solide dans la forme qui l'exprimait et servait trop à marquer les groupes à l'intérieur de la phrase, à indiquer des rapports dans des cas tels que : *je le vois* en parlant du « siège » et *je la vois* en parlant de la « chaise », pour se laisser aisément éliminer, en dépit de son manque absolu de signification. La force de résistance du genre était grande, grâce surtout à l'article *le, un* s'opposant à *la, une* ; elle a été si grande que les adjectifs comme *grand* où la distinction n'existait pas à date ancienne (on dit encore *grand mère*, et même *mère grand*, comme

grand père) ont reçu la distinction et que le féminin *grande* a été créé en face de *grand*.

Si la distinction des genres a disparu de bonne heure en iranien occidental et en arménien, c'est sans doute parce que les langues anciennes des régions où l'indo-européen a pris la forme de l'arménien et de l'iranien occidental ignoraient les distinctions de genre grammatical. On aperçoit ici l'influence du « substrat » linguistique, jointe au fait que la fin de mot qui renfermait l'expression du genre a été fortement réduite en arménien comme en iranien et que, par suite, une altération de caractère phonétique a soutenu la tendance générale et l'action des « substrats ». On doit ajouter que ces langues sont de celles où le changement du type linguistique s'est produit, en général, de la manière la plus radicale et la plus rapide.

*
* *

Au contraire, les langues slaves où les finales se sont relativement bien conservées, qui ont maintenu le type archaïque de la morphologie indo-européenne tout en innovant beaucoup dans le détail, n'ont pas seulement conservé la distinction du masculin, du féminin et du neutre dans toute son étendue. Elles lui ont donné une importance nouvelle et un caractère nouveau.

La flexion nominale des langues slaves est dominée par la distinction du masculin, du féminin et du neutre. C'est l'un des traits où l'on reconnaît non seulement l'archaïsme du slave, mais aussi la tendance archaïsante de son développement; ce trait concorde avec la longue conservation de la déclinaison des noms et avec le fait que le nombre duel a laissé dans les langues slaves modernes des traces considérables. Ce sont également des conséquences du retard des Slaves par rapport à la civilisation universelle.

Le slave n'oppose plus d'une manière générale un genre « animé », composé de masculin et de féminin, à un genre « inanimé » (neutre), comme le faisait l'indo-européen. Il a trois genres sur un pied d'égalité, le masculin, le féminin et le neutre. Il n'a guère de traits communs au masculin et au féminin, et qui opposent ces deux genres au neutre.

Du reste, les Slaves ne se sont pas bornés à conserver les anciennes notions de genre, à les grouper d'une manière nouvelle et à leur donner une importance nouvelle pour les formes de la flexion. Ils les ont étendues et développées, en créant des nuances qui n'existaient pas.

A l'intérieur du masculin singulier, il a été constitué une distinction entre un sous-genre « animé » et un sous-genre « inanimé » qui se manifeste seulement quand le nom en question est à l'accusatif, c'est-à-dire là surtout où il joue le rôle de complément direct d'un verbe : quand un nom masculin singulier désigne un être inanimé, son accusatif est identique au nominatif, tout comme s'il s'agissait d'un neutre ; au contraire quand un nom masculin de même forme désigne un être animé, son accusatif est identique au génitif, et par suite distinct du cas sujet, le nominatif. Le procédé qui consiste à exprimer de manière différente le complément direct selon qu'il est un nom d'être inanimé ou un nom d'être animé n'est pas une particularité propre au slave. On sait que, en espagnol par exemple, on « bâtit une maison », mais « on aime *à* un homme » (*a un hombre*). Dans le groupe indo-européen, cette distinction résulte toujours d'un développement nouveau, propre à telle ou telle langue.

Le développement une fois commencé s'est poursuivi. Le russe, par exemple, a étendu au pluriel l'usage qui s'était créé au singulier.

Le polonais, qui pousse souvent à l'extrême les tendances slaves communes, est allé plus loin encore. A l'ancienne distinction d'un masculin animé et d'un masculin inanimé qu'il a héritée du slave commun pour le singulier, il a superposé au pluriel la distinction d'un masculin « personnel » et d'un masculin « non personnel ». Les noms masculins désignant des personnes ont au pluriel une forme propre du nominatif (cas sujet) et offrent pour l'accusatif (complément direct) la forme du génitif, à l'imitation de ce qui arrive au singulier pour les noms d'êtres animés ; au contraire les noms masculins désignant des êtres non personnels (animaux et choses) ont au pluriel une forme commune de nominatif et d'accusatif. A côté de la dis-

tinction du masculin, du neutre et du féminin qui, comme dans l'ensemble du slave, domine la flexion nominale, le polonais offre donc au masculin singulier la distinction des sous-genres « animé » et « inanimé » (l'« animé » comprenant le « personnel ») et au masculin pluriel la distinction d'un sous-genre « personnel » et d'un sous-genre « non personnel ». Le système parvient ainsi à un maximum de complication et d'incohérence, puisque la distinction n'est pas la même au pluriel qu'au singulier, et que le masculin offre des distinctions dont il n'y a trace ni au neutre ni au féminin et puisque le cas sujet et le complément direct qui sont toujours distingués dans les féminins singuliers en *-a*, même désignant des choses, ne le sont au masculin singulier que pour les êtres animés et que, en revanche, le cas sujet et le complément direct, qui ne sont pas plus distingués au féminin pluriel qu'au neutre pluriel, même désignant des personnes, le sont dans les masculins personnels.

Ce développement du système est intervenu en polonais seulement, et durant la période historique du polonais, donc en un temps où l'influence de la civilisation occidentale s'exerçait à plein. C'est que le système de la langue a par lui-même une grande puissance et que, même si des conditions nouvelles de civilisation tendent à le détruire, il peut néanmoins continuer de se développer dans le sens où le développement a été engagé. Les langues servent à exprimer la mentalité des sujets parlants, mais chacune constitue un système fortement organisé qui s'impose à eux, qui donne à leur pensée sa forme et qui ne subit l'action de cette mentalité que d'une manière lente et partielle, seulement au fur et à mesure des occasions.

Dans l'espace d'environ quatre milliers d'années, la plupart des langues indo-européennes, organes des civilisations les plus élevées du monde, n'ont pu se débarrasser de la catégorie du genre grammatical, qui n'a aucun sens dans la plupart des cas. La même distinction du masculin et du féminin, existant en sémitique, n'y a pas non plus de sens ; et elle s'y est conservée avec plus de ténacité encore. C'est un fait qui montre à quel point un procédé linguistique une fois constitué se maintient par la seule force du système linguistique.

*
* *

Néanmoins il y a une action de la mentalité des sujets parlants. Les langues dont le développement est le plus avancé tendent soit à perdre le genre, comme il est arrivé en anglais, et aussi en arménien et en iranien occidental, soit à le réduire, comme il est arrivé dans les langues romanes qui, tout en gardant la distinction du masculin et du féminin bien vivante, ont éliminé le neutre.

Une langue d'aspect très archaïque, le lituanien, a pourtant éliminé le neutre, comme l'ont fait les langues romanes. Il peut paraître surprenant que le lituanien et le lette aient ainsi une avance sur des langues où la civilisation a agi plus vite et plus profondément. Il est probable que l'élimination du neutre a été entraînée par la tendance à distinguer le cas sujet et le complément direct dans tous les cas; le neutre, où cette distinction n'existait pas, a disparu devant le masculin, non parce que la notion de genre était contraire à la mentalité des sujets parlants, mais parce que la flexion casuelle avait gardé toute sa valeur et tendait à s'étendre, non à se restreindre. Ici encore, c'est le système linguistique dont l'influence a été décisive. C'est sans doute le retard général du développement qui a provoqué sur un point particulier un progrès non réalisé dans des langues plus avancées presque à tous égards.

Le sort du genre grammatical est ainsi l'un des faits qui illustrent le mieux, d'une part, l'action qu'exercent sur le développement de la langue la mentalité des sujets parlants et leur degré de civilisation, de l'autre, les résistances qui proviennent du système de la langue; les résistances qui résultent des formes fixées entravent l'action des faits de civilisation qui agissent sur la mentalité des sujets parlants et peuvent aller jusqu'à entraîner la langue en un sens contraire à celui où elle irait normalement. On y voit et la réalité de cette action des faits de civilisation et, d'autre part, la puissance du matériel linguistique existant.

LA CATÉGORIE DU GENRE

ET LES

CONCEPTIONS INDO-EUROPÉENNES [1]

I. — Définition du genre dans les langues indo-européennes

On enseigne d'ordinaire qu'il y avait en indo-européen trois genres : le masculin, le féminin et le neutre. Une énumération ainsi faite donne à croire que ces trois genres seraient des catégories comparables entre elles. Or, si les choses se présentent ainsi à l'époque historique, et surtout en slave, par exemple, ce n'est vrai ni pour la forme ni pour le sens en ce qui concerne l'indo-européen commun.

*
* *

Les anciens opposaient nettement le masculin et le féminin au neutre, comme on le voit par les noms qu'ont fixés les Grecs et que, suivant leur habitude, les Romains ont calqués : d'une part, ἀρρενικόν (*masculinum*) et θηλυκόν (*femininum*), de l'autre, ce qui n'est ni masculin ni féminin, οὐδέτερον (*neutrum*).

En tant qu'ils ont une valeur sémantique, les genres masculin et féminin désignent des êtres animés, de sexe mâle ou femelle, et le genre neutre désigne ce qui est inanimé. L'opposition du masculin-féminin et du neutre est donc une opposition de animé-inanimé.

Quant à la forme, le masculin-féminin (genre animé) s'oppose en effet au neutre (genre inanimé). En indo-européen, la flexion

1. Les idées exposées dans cet article ont servi de thème à des conférences faites dans les universités de Hollande et à Strasbourg.

du masculin ne se distinguait en rien de celle du féminin ; ainsi en latin, la flexion de *māter* ne se distingue pas de celle de *pater*, ou celle du féminin *fāgus* de celle du masculin *lupus* par exemple. Le caractère masculin ou féminin d'un substantif ne se reconnaissait donc en indo-européen qu'à la forme masculine ou féminine des adjectifs qui éventuellement s'y rapportaient. Ainsi, les mots latins *pater* et *lupus* sont masculins parce qu'on leur applique des adjectifs de la forme *iste*, *bonus* ; *māter*, *fāgus* sont féminins parce qu'on leur applique des adjectifs de la forme *ista*, *bona*. Sans l'accord de l'adjectif, la distinction du masculin et du féminin n'existerait pas en indo-européen.

Ce n'est que secondairement que des substantifs désignant des êtres mâles ou femelles ont reçu une forme féminine distincte de la forme masculine : quand, en grec, ἵππος désigne à la fois le « cheval » et la « jument », et admet suivant les cas un article masculin ou féminin, on est en présence de l'état des choses indo-européen ; les formes *açvā* du sanskrit, *equa* du latin, *aszvà* du lituanien, pour désigner la « jument », résultent de développements secondaires. Il se trouve que, en l'espèce, les trois langues ont abouti à des formes de même type. Mais tel n'est pas toujours le cas : en face de grec ἄρκτος « ours », masculin et féminin, le latin a *ursus*, *ursa*, mais le sanskrit *rkṣaḥ*, *rkṣī*. Si le sanskrit n'a pas, comme le grec, conservé l'usage de désigner par *áçvah* le « cheval » et la « jument », par *rkṣah* l' « ours » et l' « ourse », c'est qu'il a, d'une manière générale, perdu l'emploi des thèmes en *-a-* bref au féminin, tandis que les substantifs grecs correspondants en *-o-* ont gardé la possibilité d'être féminins. Mais, là où la forme permettait de désigner par un même mot le mâle et la femelle, le sanskrit a, comme le grec, maintenu l'usage indo-européen : le sanskrit *gaúḥ* désigne le « bœuf » et la « vache », tout comme grec βοῦς et latin *bōs*.

Dans les adjectifs qui étaient seuls à caractériser l'opposition du masculin et du féminin, la distinction se marquait par une différence, non de flexion, mais de thème : un thème masculin **néwo-* (sanskrit *návaḥ*, gr. νέϝος, lat. *nouus*, v. slave *novŭ*) s'oppose à un thème féminin * *néwā-* (skr. *náva*, gr. νέϝα, lat. *noua*, v. sl.

nova). Le thème du masculin est le principal : il sert aussi pour le neutre (skr. *návam*, gr. νέϝον, lat. *nouum*, sl. *novo*) ; c'est la forme normale. Dans la plupart des types, la forme féminine est dérivée de la forme masculine, au moyen d'un suffixe : ainsi, dans les participes présents, on a au masculin (et au neutre) un thème en *-ont-/-nt-* et au féminin *-ntī*, *-ntyā-* ; soit en sanskrit, au masculin, *uçán* « voulant », génitif *uçátaḥ*, la forme féminine correspondante est *uçat-ī́* ; en grec, ϝεκών a pour féminin ϝέκουσα (ϝεκαθθα du crétois répond bien au type sanskrit). Le genre féminin apparaît ainsi comme un sous-genre à l'intérieur du genre animé.

Quand on veut désigner l'« homme », en général, le genre est toujours masculin, si divers que soient les mots, ainsi *homō* en latin ou ἄνθρωπος en grec. Et un adjectif qui se rapporte à deux substantifs, désignant des personnes l'un masculin, l'autre féminin, se met au masculin, ainsi chez Homère, ο 162 :

οἱ δ' ἰύζοντες ἕποντο
ἄνέρες ἠδὲ γυναῖκες.

On voit par là que, en même temps qu'il désigne spécialement le mâle, le masculin est le genre commun, et que le féminin en est seulement une différenciation. Là où il est question d'« oies » en général, Homère traite χῆνες comme un masculin, τ 552-553, et le féminin n'apparaît que là où il est question d'une oie (évidemment femelle) isolée, ο 161, 174.

Si certains animaux ont des noms féminins, sans considération de sexe, ce ne sont que de petits animaux, surtout des insectes, ainsi la « guêpe », en latin *uespa*, en vieux haut allemand *wafsa*, en lituanien *vapsà*.

*
* *

Au genre animé, marqué par le masculin, avec une différenciation éventuelle pour le cas particulier du féminin, s'oppose le genre inanimé, le « neutre ».

La distinction du masculin (-féminin) et du neutre se marque au moyen de la flexion qui est nettement différente pour les deux genres ; mais elle n'est faite qu'au nominatif, à l'accusatif et au vocatif. Le genre animé a des formes distinctes pour ces trois cas, comme le sens le demande en effet : le vocatif est un cas important pour les noms de personnes et d'êtres animés en général ; la distinction du « sujet », qui agit, et de l'« objet », qui subit l'action, est essentielle pour les noms d'êtres animés : il importe de savoir si *Paul bat Pierre* ou si *Pierre bat Paul* ; les langues où les formes du nominatif et de l'accusatif se sont confondues, mais où subsistent des distinctions casuelles, tendent par suite à distinguer par quelque moyen la forme du sujet et celle du complément direct pour les noms d'êtres animés, surtout pour les noms de personnes : le slave, où dans le type le plus courant du masculin, le nominatif et l'accusatif singuliers se sont confondus, a substitué à l'accusatif une autre forme casuelle, celle du génitif, quand il s'agit d'êtres animés ; comme, en arménien moderne, la confusion des formes du nominatif et de l'accusatif est complète, beaucoup de parlers arméniens actuels expriment par le datif le complément direct quand ce complément est un nom d'être animé. Mais ainsi qu'on le voit par le slave et par les parlers arméniens, la confusion des formes du nominatif et de l'accusatif est facilement tolérée pour les noms d'êtres inanimés, les noms de choses ou les abstraits. Les noms de genre inanimé n'ont bien entendu pas besoin de vocatif.

Tous les accusatifs des noms de genre animé (masculins ou féminins) sont caractérisés par une même désinence indo-européenne, au singulier par une nasale finale (type latin *lupum* ou grec λύκον), au pluriel par **-ns* (ainsi *stans* en vieux prussien, τονς en grec crétois), cette finale étant sujette à de grandes altérations phonétiques qu'il est inutile d'indiquer ici. De tous les cas de la déclinaison indo-européenne, c'est celui qui est marqué de la manière la plus claire et la plus uniforme. Le nominatif singulier des noms du genre animé est marqué soit par une particularité du vocalisme (en général le degré long, comme dans grec πατήρ, ἡγεμών), soit par une désinence *-s* (comme dans grec

λύκος, latin *lupus*, lituanien *vilkas*, etc.). M. Schuchardt a fait remarquer combien cet emploi d'une désinence *-s* au cas sujet est singulière : dans la plupart des langues, le cas sujet est exprimé par la forme même du nom, sans désinence. Or, le nominatif-accusatif neutre est précisément la forme du mot sans désinence dans la plupart des neutres indo-européens, ainsi pour le mot qui signifie « miel » et « hydromel », sanskrit *mádhu*, grec μέθυ, vieux prussien *meddo*, etc.

Chose curieuse, cette forme de nominatif-accusatif neutre semble souvent échapper à toute flexion. La flexion des cas autres que le nominatif-accusatif est fréquemment empruntée à un thème autre que celui qui fournit le nominatif-accusatif ; c'est ce qui arrive notamment dans le type connu que représente par exemple le nom du « foie » : thème en *-r-* au nominatif : sanskrit *yákṛt*, grec ἧπαρ, latin *iecur* ; thème en *-n-* aux autres cas : sanskrit *yaknáḥ*, grec ἥπατος, latin *iecinoris*. La forme qui tient la place du nominatif-accusatif pluriel, lequel n'existe pas, était une forme de nominatif-accusatif neutre de thème en **-ā-*, dont il n'existe aucune forme casuelle ; sanskrit *yugā* « jougs », grec ζυγά, latin *iuga*. A date historique, la forme de nominatif-accusatif pluriel neutre tend à entrer dans le type général de la flexion ; mais en indo-européen, elle était tout à fait aberrante.

Ces faits — et l'on ne marque ici que les principaux — montrent que la distinction entre le genre animé (masculin-féminin) et le genre inanimé (neutre) et la distinction entre le masculin et le féminin sont hétérogènes.

II. — Les noms de l'« eau » et du « feu » [1]

Pour désigner l'« eau » et le « feu », on observe en indo-européen deux séries de noms, les uns de genre neutre (inanimé), les autres de genre masculin ou féminin (animé). Ces deux

1. Une étude plus technique de ces noms se trouve dans les *Mémoires de la Société de Linguistique*, XXI, p. 249 et suiv.

types de noms traduisent évidemment des conceptions différentes.

Pour l'« eau », le sanskrit védique a deux dénominations dont il est facile de déterminer la valeur. L'une est de genre féminin, c'est le thème *ăp-*, qui se retrouve en iranien et qui a subsisté, par exemple, dans le persan *āb* ; ce mot s'emploie surtout au pluriel : *ăpaḥ* « eaux » ; il désigne les « eaux », en tant qu'elles sont considérées comme des êtres qui agissent et, par suite, comme des forces naturelles de caractère religieux ; ce sont les « eaux divines », *ăpo devíḥ*, *ăpo divyáḥ*, les « eaux mères », *ăpo mātáraḥ* ; le dieu Feu, Agni, est leur fils en tant que « descendant des eaux » *apắṃ nápāt*. Au contraire, le nom neutre, *udakám* au nominatif-accusatif singulier, *udán-* aux autres cas (génitif *udnáḥ*, locatif *udáni*), désigne l'« eau » considérée comme une chose, c'est-à-dire comme la conçoit un moderne. Dans le Rg-Veda, V, 45, 10, on lit :

udnā́ ná nā́vam anayanta dhā́rā
āçṛṇvatír ăpo árvāg atiṣṭhan

« comme les sages ont conduit un navire sur l'eau, les eaux dociles se sont arrêtées ». Le navire flotte sur l'« eau », considérée comme une chose, et nommée du nom neutre *udán-* ; mais dès que les eaux sont considérées comme des êtres agissants, qu'il s'agit de leur docilité, de leur arrêt, le poète se sert du nom féminin, *ăpaḥ*, au pluriel. Cet exemple met en pleine évidence la valeur des deux termes par lesquels les poètes védiques désignent l'« eau ».

Donc, suivant qu'il était considéré comme une chose ou comme un être susceptible d'action, un même corps, l'eau, pouvait recevoir deux noms, l'un de genre neutre, l'autre de genre animé, en l'espèce, de genre féminin. C'est dire que, à une conception de ce corps comme un objet matériel, conception analogue à la conception d'un civilisé actuel, se juxtaposait une conception tout autre, suivant laquelle les objets et les phénomènes naturels sont les manifestations de forces internes, ana-

logues à celles qui meuvent les animaux et les hommes ; ces forces étaient associées à des notions religieuses, et on leur attribuait quelque chose de « divin ».

Du reste, l'importance des catégories de genre dans la langue remonte évidemment à un temps — sans doute en partie antérieur à celui où s'est fixé l'indo-européen commun — où ces conceptions de demi-civilisés qui voient partout des forces actives, analogues à celles des êtres animés, prévalaient encore. Le livre fondamental de M. Lévy-Bruhl sur *les Fonctions mentales dans les sociétés inférieures* montre la place de ces conceptions dans la pensée humaine ; et M. de Josselin de Jong a étudié le problème au point de vue linguistique dans son *De waarderingsonderscheiding van « levend » en « levenlos »* (Leide, 1913). Il est curieux que, en général, les noms d'arbres indo-européens soient de genre féminin, et les noms de fruits de genre neutre : l'arbre était considéré comme une sorte de femelle qui produit des fruits ; cette opposition s'observe parfois en slave où le nom *(j)ablŭko* de la « pomme » est généralement neutre et le nom *(j)ablanĭ* du « pommier » féminin ; elle est régulière en grec et en latin ; ainsi le latin a *pirus* (féminin) « poirier » et *pirum* (neutre) « poire». Cette opposition est si naturelle qu'elle se retrouve en algonquin, comme l'a signalé M. de Josselin de Jong, *l. c.*, p. 134 et suiv.

Le nom neutre de l' « eau », traduisant la conception toute matérielle et non religieuse, tel qu'on le rencontre dans skr. *udakám, udnáh*, se retrouve en grec sous la forme ὕδωρ, ὕδατος, et en ombrien sous la forme *utur, une* (ablatif), en germanique sous la forme de *wato* en gotique, *vatn* en vieil islandais et de *waeter* en vieil anglais, de *wazzar* en vieux haut allemand. C'est un de ces noms neutres indo-européens, de structure assez compliquée, dont le nominatif-accusatif singulier était en *-r*, et les autres cas en *-n-*, et dont l'élément radical offrait des alternances vocaliques corrélatives. Par suite de la complication extrême de la flexion de cet ancien neutre, le mot a subi partout des altérations. Le grec et l'ombrien, qui gardent l'opposition de **-r* du nominatif-accusatif et de **-en-* des autres cas, ont généralisé le vocalisme radical zéro, type **ud-*. Au contraire, le germanique, où chaque dialecte

a généralisé soit le type en **-r-* soit le type en **-en-*, a généralisé un vocalisme constant *-o-*, soit **wod-*, germ. **wat-*.

Le slave a aussi généralisé **wod-*, mais en faisant passer le mot par dérivation au genre féminin : *voda*. On obtient ainsi un mot de genre animé, dont la valeur est comparable à celle de l'indo-iranien **ăp-*.

Le latin a aussi un mot de genre animé, et aussi féminin : *aqua*, qui se retrouve en germanique, gotique *ahwa*, v. islandais *ǫ*, v. haut allemand *aha* etc. Le caractère religieux du mot serait bien net si l'on admettait l'hypothèse, soutenue par M. Noreen, que le nom de dieu nordique de la mer *ægeir* appartenait à cette racine, et qu'il y aurait ici une alternance *ē/ǫ*.

Mettant ainsi en évidence le caractère « animé » et, par suite, divin de l'eau, le latin est allé plus loin que le germanique ; il n'a pas seulement *aqua* ; il a aussi fait passer le vieux nom neutre au type « animé » en ajoutant le suffixe **-ā-*, comme le slave. Mais il a fait plus. Le nom skr. *udakám* appartient à une racine qui fournit un verbe ; or, le présent sanskrit de cette racine est du type à infixe nasal, soit *unátti* « il mouille, il jaillit », 3e plur. *undánti* ; ce présent, s'il était conservé, serait en latin de la forme **undō*, dont la nasale infixée a été transportée dans le substantif ; ceci montre bien que *unda* désigne l'eau comme quelque chose qui agit, comme un être animé. Du reste *unda* se distingue de *aqua* en ce que ce mot désigne plutôt l'eau en mouvement, la vague, l'eau jaillissante, que l'eau considérée comme un objet, comme un élément ; aussi est-ce souvent un mot poétique ; et le sens « actif » a persisté dans *onde* du français, *onda* de l'italien et de l'espagnol. Ailleurs encore l'*n* du présent à infixe nasal a pénétré dans le nom de l'« eau » ; on observe le fait dans les langues baltiques : en vieux prussien, où le nom dérivé *wundan* est neutre, et en lituanien, où sans doute le masculin *vandů* repose sur un ancien neutre (le lituanien a perdu le genre neutre, et les neutres y sont représentés par des masculins).

Le nom neutre, gr. ὕδωρ, etc., de l'« eau » n'est pas le seul connu. Il y a en sanskrit un mot *vâr* (forme védique ; classique *vâri*, gén. *vâriṇaḥ*), et en « tokharien A », *wär* ; on en rapproche

vieil anglais *waer* (neutre) « mer » et *ear* (aussi neutre), qui a le même sens, et grec οὖρον « urine ». Le vieux prussien *wurs* qui, dans le vocabulaire d'Elbing, traduit *tych* (*teich*), est masculin ; il en est de même du mot avestique *vār-* qui désigne la « pluie » : on voit qu'un même mot indo-iranien,* *wār-*, est neutre ou masculin suivant qu'il désigne l'eau en repos ou l'eau en mouvement.

L'arménien a *jur* (génitif *jroy*), qui paraît reposer sur **yuro-*, neutre, qui rappelle lituanien *júrės* « mer », et qui serait à lituanien *jaura* « endroit marécageux », ce que vieux prussien *wurs* est à vieil anglais *ear*, gr. οὖρον.

Les mêmes langues où le nom de l'« eau » est de genre « animé » et où l'« eau » est, pour ainsi dire, personnifiée et susceptible d'être considérée comme divine, ont aussi pour le « feu » un nom de genre animé, masculin, et le « feu » y est un être divin. Ceci est surtout visible dans l'Inde, où le nom sanskrit du « feu », *agníḥ*, est celui d'une des principales divinités védiques. Le mot *agníḥ* a des correspondants en latin (*ignis*), où les noms de l'« eau » sont le nom féminin *aqua* et le nom féminin *unda*, en slave (*ognĭ*, *ognji*) où le nom de l'« eau » est le nom féminin *voda*, en letto-lituanien (lituanien *ugnìs*, lette *uguns*) où le nom de l'eau (lit. *vandů*) se rattache, on l'a vu, au type animé, au moins secondairement. L'iranien, où les termes religieux sont en partie autres que dans l'Inde, offre un mot tout différent, mais également masculin, avestique *ātarš*, dont des formes arméniennes et slaves attestent l'antiquité.

Au contraire, les langues où le nom de l'« eau » est de genre neutre, ont des noms neutres du « feu ». Le plus clair est celui qu'on trouve dans πῦρ du grec, *pir* de l'ombrien, *fiur* du vieux haut allemand, *hur* de l'arménien. Un autre tout différent, et moins clair, est sûrement aussi ancien : gotique *fon* (génitif *funins*), vieux prussien *panno* (ancien *paňů*).

Il résulte de là que l'indo-européen a eu pour l'« eau » et le « feu » deux types de noms, les uns de genre animé, masculins

pour le « feu », féminins (ou féminisés) pour l'« eau », les autres du genre inanimé (neutre). En indo-européen, les deux types coexistaient, et pour l'« eau », le sanskrit et le germanique ont encore les deux côte à côte. Mais, la plupart du temps, il n'a subsisté dans une langue donnée que l'un ou l'autre. Il est curieux que le choix de dialectes parfois très voisins soit différent ; ainsi le latin n'a que les formes du genre animé, *aqua* et *unda* pour l'« eau », *ignis* pour le « feu », et l'ombrien (et sans doute l'osque) que les formes du genre inanimé : *utur* pour l'« eau », *pir* pour le « feu ». Ceci paraît montrer que la limitation à l'un des deux types est chose relativement récente.

Le fait de choisir soit le type « animé », soit le type « inanimé », caractérise les langues. Là où, comme dans l'Inde ou à Rome, prévalent les préoccupations religieuses, les formes de genre « animé » tendent aussi à prévaloir. Là où, au contraire, comme en Grèce, les points de vue profanes dominent et où la pensée est toute « laïque », les formes de genre inanimé ont seules persisté. La prédominance de ὕδωρ et de πῦρ est une des marques du fait que les Grecs voyaient les choses d'une manière profane et matérielle. Leurs conceptions sont déjà modernes, et les vieilles conceptions animistes n'existent plus chez eux qu'à l'état de traces.

III. — De quelques noms d'astres

Le nom de genre animé et le nom de genre inanimé ne diffèrent souvent que par la formation. Le nom de l'« eau » en a fourni déjà un exemple : en face du neutre *udakám*, *udnáḥ* du sanskrit, ὕδωρ, ὕδατος du grec, *utur*, *une* de l'ombrien on a les dérivés féminins, *voda* du slave, *unda* du latin.

Le nom du « soleil » fournit un exemple analogue.

Il y en a une vieille forme neutre, offrant la flexion caractéristique avec nominatif-accusatif appartenant à un thème autre que celui qui fournit les autres cas. La forme est surtout claire en iranien : l'Avesta a au nominatif-accusatif *hvarə* (c'est-à-dire *huvar*) et au génitif singulier, dans les *gāthās*, *xvəng* (représentant

un ancien indo-iranien **swans*). Le sanskrit n'a gardé que la forme du nominatif-accusatif, védique *s(ú)var*, sur laquelle il a été refait un génitif nouveau, *súraḥ*. Une forme neutre reposant sur le nominatif-accusatif se trouve aussi dans le gotique *sauil*, et un dérivé de forme compliquée dans *slŭnĭce* du vieux slave, qui repose sur une forme slave commune déjà neutre.

Mais le « soleil » peut être considéré comme un être agissant et susceptible d'être divinisé. Aussi a-t-on, à côté des formes neutres, des formes de genre animé, masculin ou féminin : latin *sōl* (acc. *sōlem*) est masculin ; le vieil islandais *sól* est féminin, ainsi que le dérivé gotique *sunno*, vieil islandais *sunna*, vieux haut allemand *suṇna*, etc. de la forme des cas en *-n-*. Le sanskrit a des dérivés masculins, l'un en *-a-*, *súraḥ*, l'autre en *-iya-*, *sūr(i)yaḥ*. Le grec a une forme pareille, mais indépendante, à en juger par le vocalisme radical : *hāϝελιος* (crétois ἀβέλιος, homérique ἠέλιος, attique ἥλιος). Dans les langues baltiques, le nom est un dérivé féminin de la forme du nominatif-accusatif : lituanien *sáulė*, vieux prussien *sâule*. C'est que le « soleil » est tenu pour un être divin. Dans l'Avesta, le nom du « soleil », quoique neutre, est celui d'un être « divin », ce qui est assez singulier. Plus logiques, les Grecs, qui ont conservé la conception du « soleil » divinisé, ont donné à son nom une forme de genre masculin.

A ce cas s'opposent ceux des « étoiles » et de la « lune » dont le nom est toujours de genre animé.

Le nom grec ἀστήρ de l'« étoile » est masculin ; il en est de même de *star-* en indo-iranien (et de *tārāḥ*, nominatif pluriel en sanskrit) ; le dérivé qu'on observe en germanique occidental, *sterno* en vieux haut allemand, *steorra* en vieil anglais, est aussi masculin. D'autres dérivés, *stēlla* (de **stēr-lā*) en latin, *stairno* en gotique et *stiarna* en vieil islandais, *seren* en gallois, sont féminins. Le neutre grec ἄστρον a sans doute été fait secondairement sur le collectif ἄστρα, qui est la forme la plus ordinaire, et la seule attestée chez Homère. Il est curieux de voir ici comment les étoiles, dont la répartition varie suivant les saisons, ont un nom de genre « animé », et le neutre n'intervient que là où l'on envisage des ensembles.

La « lune », astre non seulement mobile, mais changeant, a toujours un nom de genre « animé » : en attique μήν, en gotique *mena*, etc., toujours masculins; ce nom ne désigne pas seulement la lune, mais aussi la « lunaison », le « mois ». Il en existe un dérivé féminin en grec, dorien μήνᾱ, ionien-attique μήνη. Là où la « lune » est désignée par un adjectif pris substantivement, comme dans latin *lūna*, vieux slave *luna*, ou dans grec σελήνη— tous mots signifiant « brillant » —, le genre est féminin (on ne saurait rien dire de l'arménien *lusin*, dont les formes mêmes sont flottantes).

IV. — Les noms du « songe »

Le nom du « sommeil », ὕπνος en grec, *somnus* en latin, etc., est masculin, parce que le « sommeil » est une force puissante qui soumet les hommes à sa volonté.

Au contraire, l'objet qui apparaît dans le sommeil, le « songe », est partout du neutre : ἐνύπνιον en grec, *somnium* en latin, *sŭnĭje* en slave, *svápn(i)yam* en sanskrit.

Le grec a un autre mot, dont un correspondant se trouve en arménien, et qui mérite d'être observé de près. Ce mot a des formes neutres, qui sont celles du primitif, et des formes masculines, manifestement dérivées.

La forme primitive du mot est de type spécifiquement neutre : thème en *-r* au nominatif-accusatif singulier, en *-n-* aux autres cas : ὄναρ est conservé au nominatif-accusatif singulier ; au génitif on attend *ονατος (type ἦπαρ, ἥπατος en face de *yákṛt* « foie », génitif *yaknáḥ* du sanskrit) ; mais le grec évite la suite de *-*nn-* ; à côté de θέναρ on n'a que θέναρος et non le *θενατος attendu ; à côté de ὄναρ, on a, dès l'époque homérique, ὀνείρατος, avec -ειρ- emprunté au dérivé ὄνειρος, ὄνειρον. La vieille forme ὄναρ, ὀνείρατος désigne le « songe » neutre, sans indiquer que le « songe » soit personnifié.

La forme dérivée, de genre masculin, ὄνειρος (ὄνοιρος en éolien) se prête au contraire à désigner le « songe » personnifié. Quand, au chant B de l'Iliade, Zeus appelle le « Songe », l'interpelle,

l'envoie à Agamemnon, le nom est toujours ὄνειρος, jamais ὄναρ. Quand le songe est considéré comme actif, c'est à ὄνειρος que recourt le poète, ainsi dans l'Odyssée :

λ 206 τρὶς μὲν ἐφωρμήθην, ἐλέειν τέ με θυμὸς ἀνώγει,
τρὶς δέ μοι ἐκ χειρῶν σκιὴ (ϝ)είκελον ἢ καὶ ὀνείρῳ
ἔπτατο.

La valeur active de ὄνειρος n'est pas toujours sensible ; par exemple elle l'est peu au chant τ de l'Odyssée, dans le récit du songe de Pénélope.

Ce n'est pas à dire d'ailleurs que la forme dérivée soit toujours masculine ; un neutre ὄνειρον se lit déjà chez Homère et se retrouve par la suite ; et, à côté de ὄναρ, le crétois a ἄναιρον, glosé par ὄνειρον chez Hesychius, et qui est sans doute aussi neutre.

La différence entre ὄναρ, ὄνειρον (en crétois ἄναρ, ἄναιρον) et ὄνειρος traduit une différence de conceptions qu'on entrevoit encore en grec, mais qui, dès le début de l'époque historique, tend à s'effacer. Par malheur, le mot ne se retrouve, hors du grec, qu'en arménien, où les distinctions de genre grammatical ont disparu avant l'époque historique. On peut seulement constater que l'arménien *anurǰ* est un dérivé, assez analogue sans doute à ὄνειρος (ὄνειρον), ὄνοιρος, ἄναιρον ; le dérivé repose sur une ancienne forme de nominatif-accusatif neutre **anōr*, qui est à crétois ἄναρ ce que, en grec, τέκμωρ est à τέκμαρ.

V. — Les noms du « jour » et de la « nuit »

De même que pour le « feu » et l'« eau », il y a, pour le « jour », des expressions de genre « animé » et de genre « inanimé ». Le contraste entre les unes et les autres est frappant, bien que peu de langues les conservent à la fois.

Quoiqu'il ait perdu le genre grammatical, l'arménien a gardé un représentant de l'ancien nom masculin, à savoir *tiw* « jour » (en tant qu'opposé à la « nuit »), et un représentant d'un des anciens neutres, *awr*, (« journée » espace de vingt-quatre heures),

et aussi « jour » (distinct de la nuit). En sanskrit, le masculin *dyaúḥ* n'apparaît que dans le sens de « jour » et surtout dans des formes adverbiales, notamment dans *divā* « de jour ». Le nom ordinaire du « jour », même en tant qu'il s'oppose à la nuit, est *áhar*, gén. *ahnáḥ*, forme de type essentiellement neutre ; ce nom neutre se retrouve en iranien, dans des formes avestiques telles que le locatif *asni* ou le génitif pluriel *asnąm* (le perse remplace ce mot par un abstrait qui signifie « lumière », vieux perse *rauča*, persan *rōz*).

L'histoire du nom masculin est remarquable. Il s'agit d'un mot qui désigne la « lumière », en tant qu'elle est active, à la fois le « ciel lumineux » et le « jour ». Divinisée, cette notion se présente comme le dieu principal des Grecs, Ζεύς, Δι(ϝ)ός, et des Romains, *Iuppiter*, *Iouis*, ou des Ombriens : *Iupater*, *Iuve*. Dans le véda, *dyaúḥ* (génitif *diváḥ*, locatif *dyávi*) est aussi un personnage divin, mais qui n'est pas au premier plan, comme l'étaient Zeus et Jupiter. La flexion de ce mot offre des particularités qui n'avaient sans doute rien d'anomal en indo-européen, mais qui ne se sont pas maintenues en général, tandis qu'elles subsistaient dans ce mot important. Le védique qui a *dyaúḥ* (ou *diyaúḥ*) au nominatif, *dyắm* (ou *diyắm*) à l'accusatif, *diváḥ* au génitif-ablatif, *dyávi* au locatif, en donne une idée ; le latin a gardé deux séries de formes, mais en tient les flexions différenciées : le type correspondant à celui de sanskrit *diváḥ*, *dyávi* a été réservé au nom du dieu, *Iouis*, *Iouī*, *Ioue*, et l'on a tiré de là un accusatif *Iouem* (l'addition de *pater* dans la forme de nominatif-accusatif *Iuppiter* sert à éviter le monosyllabisme qui a persisté dans les formes grecques correspondantes Ζεύς, Ζεῦ) ; le type correspondant à celui de l'accusatif védique *d(i)yắm* s'est maintenu au sens de « jour », *diem*, et l'on en a tiré toute une flexion sur un modèle latin existant, *diēs*, etc. Originairement masculin, le mot latin *diēs* est devenu en partie féminin, sous l'influence des autres noms en *-iēs* d'une part, du nom de la « nuit » de l'autre. Comme pour l'« eau » et le « feu », le latin a généralisé un nom indiquant les choses conçues comme animées. Et c'est d'autant plus remarquable que ce nom est, pour ainsi dire, une création du latin, partie d'une seule forme indo-européenne.

A côté de ce mot, le slave en présente un autre, qui a la même racine, mais une formation différente, à savoir un thème en *-n-*, représenté en vieux slave par *dĭnĭ*. Ce mot n'a en slave que le sens de « jour » ; et les formes correspondantes d'autres langues n'ont aussi que ce sens ; on a, par exemple, en sanskrit *madhyáṃ-dinaḥ* (masculin) « midi », latin *nūn-dinae*, littéralement « les neuf jours », et surtout en baltique le mot féminin lituanien *dẽnà*, vieux prussien *deinan* (à l'accusatif).

Le mot de genre neutre de l'indo-iranien, que représente *áhar*, *ahnáḥ* du sanskrit, ne se retrouve nulle part ailleurs. Mais il y a un autre mot, aussi neutre, en grec, c'est-à-dire dans une langue où ont prévalu les dénominations neutres de l'« eau » et du « feu » : homérique ἦμαρ, ἤματος ; l'arménien *awr*, génitif *awur* en est le correspondant, avec une finale ancienne *-ōr* (cf. τέκμαρ, τέκμωρ). Toutefois, cet ancien neutre a été remplacé par un dérivé féminin, comme il est arrivé dans le type de slave *voda* « eau », etc. ; soit ἀμέρᾱ (en attique ἡμέρᾱ, avec un esprit rude secondaire). Si la forme ἦμαρ, ἤματος domine de beaucoup chez Homère, c'est que ἡμέρη entrait malaisément dans l'hexamètre au nominatif, et pas du tout aux autres cas. Dans les rares exemples où figure ἡμέρη, on aperçoit la valeur active du mot, notamment dans ce vers de l'Iliade, Θ 541 = N 828 :

ὡς νῦν ἡμέρη ἥδε κακὸν φέρει Ἀργείοισιν.

Il est du reste possible que l'opposition du mot « nuit » ait contribué à faire généraliser la forme dérivée féminine.

En effet, la « nuit » dont le caractère religieux est beaucoup plus vivement senti que ne l'est celui du « jour », parce qu'elle a quelque chose de plus mystérieux, a partout un nom féminin : le nom indo-européen attesté par grec νύξ, latin *nox*, gotique *nahts*, lituanien *naktìs*, etc. est partout féminin ; le mot indo-iranien, qui n'a pas de correspondant ailleurs, sanskrit *kṣap-*, avestique *xšap-* est aussi féminin (on ne saurait faire état des neutres *xšapar-*, *xšapan-*, formes avestiques secondaires, manifestement faites d'après les noms du « jour ». En sanskrit *rātrī* « nuit » est également féminin, et a une forme du type féminin.

De même que le nom « animé » de « l'eau » est souvent au pluriel en indo-iranien, et le nom « inanimé » (neutre) au singulier, il est curieux qu'Homère ait souvent pour « nuit et jour » la formule νύκτας τε καὶ ἦμαρ, littéralement « les nuits (au féminin) et le jour (au neutre) ».

VI. — Les noms de parties du corps

Les noms de parties du corps demanderaient un examen assez long. Ils posent beaucoup de menus problèmes. Mais, sans entrer dans ce détail, qui serait infini, on voit que les organes actifs avaient en indo-européen des noms masculins ou féminins, et les organes considérés comme non agissant, des noms de genre neutre.

Le vieux nom du « pied », *pât* en sanskrit, πούς en grec, *pēs* en latin, *fotus* en gotique, est masculin.

Les noms de la « main » sont variés ; sauf le nom indo-iranien, sanskrit *hástaḥ*, avestique *zastō*, qui est masculin, ils sont de genre féminin ; tel est le cas de χείρ en grec, de *manus* en latin, de *handus* en gotique ; la « main » est en général nommée au féminin évidemment parce qu'elle sert à « recevoir » des objets.

Le nom de l'« avant-bras » (d'où parfois le nom du « bras » pour lequel il n'y a pas de mot indo-europen) et de la « jointure du bras », *bāhuḥ* en sanskrit, πῆχυς (πῆχυς) en grec, *bógr* en vieil islandais, est masculin.

Le nom de la « langue » dont les formes offrent des différences si étranges d'une langue à l'autre, est partout féminin : c'est que la « langue » est particulièrement active.

Au contraire, les organes immobiles ont des noms de genre neutre. C'est le cas des organes internes, notamment le « foie », grec ἧπαρ, ἥπατος, latin *iecur*, etc. ; ou, chose plus singulière, le nom du « cœur », grec κῆρ (à côté d'un dérivé καρδία, il est vrai ; ce passage au type « animé » est exceptionnel pour le nom du « cœur »), latin *cor*, gotique *hairto*, etc. Le nom de l'« os », sanskrit *ásthi* (génitif *asthnáḥ*), grec ὀστέον, latin *os* (*ossis*) et *ossu* (*ossua*) est neutre ; arménien *oskr* repose sans doute aussi sur une forme neutre. Parmi les noms d'organes immobiles, il en

est un qui est de genre « animé » (masculin ; rarement féminin); c'est celui de la « rate ». Mais les divergences singulières qu'offrent les noms de cet organe, d'une langue à l'autre, montrent que des idées religieuses y étaient attachées. Et l'on a ainsi une confirmation du principe.

Tous les cas ne sont pas aussi clairs. On peut imaginer pourquoi le nom du « coude » est de genre animé (masculin : sanskrit *aratníḥ*, vieux slave *lakŭtĭ*, ou féminin : grec ὠλένη, latin *ulna*, vieux haut allemand *elina*) et le non du « genou » neutre (sanskrit *jánu*, grec γόνυ, etc.); mais on n'aurait sans doute pas prévu cette différence.

VII. — Noms d'action et abstraits

Les noms d'action en indo-européen sont les uns masculins, les autres féminins, sans qu'on voie la raison de cette différence; mais tous sont de genre « animé » et non pas « neutre ». Qu'il s'agisse du type de latin *uōx*, de celui de grec λόγος, de celui de φυγή, de celui de grec δόσις ou de latin *datiō*, de celui du latin *ad-uentus*, tous les noms de cette sorte sont de genre « animé ».

Les noms neutres indiquent des choses ou des abstractions, non des actes ; ainsi le grec oppose le neutre ἀνά-θημα « offrande déposée » au féminin ἀνά-θεσις « fait de déposer une offrande ». Un neutre en *-es-* désigne une chose établie : le sanskrit *śrávaḥ* et le grec κλέ(ϝ)ος « réputation », l'avestique *sravō* et le slave *slovo* (génitif *slovese*) « parole ».

Ce n'est pas un accident que le grec ait en général le neutre τέρμα pour signifier « borne, limite », et le latin le masculin *terminus* : on sait que, à Rome, *terminus* a un caractère divin, que le grec τέρμα ne possède pas. Le grec a du reste aussi, chez les poètes seulement, et avec une nuance de sens franchement religieuse, le masculin τέρμων, qui est le synonyme de τέρμα, à la nuance religieuse près. Il y a d'ailleurs trace en latin aussi du neutre *termen*, qui n'a pas prévalu dans l'usage.

Les noms d'instruments, dont le suffixe offre des formes diverses, sont neutres : on a donc, toujours au neutre, en grec

ἄροτρον « charrue », en latin *arâtrum*, en slave *ralo* (ancien *radlo*). Mais le même suffixe peut fournir un nom masculin si ce nom est celui d'une notion susceptible d'être considérée comme animée, et comme divine, ainsi *vĕtrŭ* « vent » (d'une racine **wē-* « souffler ») en slave.

VIII. — Conclusion

Ces principes une fois posés, il n'est guère de substantif indo-européen dont le genre ou « animé » (masculin ou féminin) ou « inanimé » (neutre), ou flottant ne se laisse aisément justifier. Inintelligible à l'époque historique où elle n'est qu'une survivance, la catégorie du genre avait sa pleine valeur en indo-européen, c'est-à-dire dans la langue d'un peuple qui opposait d'une manière systématique et constante l'animé à l'inanimé.

La différence entre le masculin et le féminin ne laisse, au contraire, presque jamais remonter à une signification définie, sauf les cas, peu nombreux en somme, où elle sert à marquer l'opposition du « mâle » et de la « femelle ». Dans la mesure où le féminin s'oppose au masculin, il apparaît comme une différenciation : le masculin indique d'une part, le type générique, et de l'autre, en particulier, le sens mâle ; ceci s'exprime par le fait morphologique que la forme du féminin est dérivée de celle du masculin : le féminin *gurv-ī* du masculin sanskrit *gurú-ḥ* « lourd » représente la forme du masculin, plus un suffixe secondaire. Et ainsi presque toujours.

Il semble que, le plus souvent, la différence du masculin et du féminin soit affaire de pure forme : les thèmes en *-o-* de genre « animé » sont tous masculins, pour autant que le sens n'impose pas le genre féminin, ce qui arrive notamment dans le nom de la « bru », grec νυός, etc. ; et dans les noms d'arbres tels que gr. φηγός ; les thèmes en *-ā-* sont tous féminins, pour autant qu'ils ne désignent pas expressément des mâles, comme lat. *scrība* et *aurīga* ; et ceci se conçoit bien puisque dans les adjectifs, un thème en *-o-* caractérise le masculin, et un thème en *-ā-* le féminin : grec νέ(ϝ)ος : νέϝᾱ « neuf : neuve ». De même, les thèmes en *-i-* de genre

« animé » sont en général féminins, et les thèmes en *-u-* de genre « animé », masculins, ainsi que les thèmes en *-n-*. Les noms radicaux sans aucun suffixe sont du féminin pour peu que le sens n'exige pas le masculin; ainsi latin *uōx* et sanskrit *vāk* « voix »; grec στύξ, φλόξ, etc. On n'aperçoit souvent pas pourquoi tel ou tel type de mots, tel ou tel mot en particulier, appelle un adjectif de forme masculine ou de forme féminine. Pour autant qu'on aperçoit les raisons, elles tiennent à des conceptions qui se relient à celles qui conditionnent le genre animé ou inanimé des mots : le « ciel » d'où vient la pluie fécondante est du masculin, la « terre », qui est fécondée, est du féminin; le « pied » est du masculin, la « main, qui reçoit, est du féminin.

Ainsi la question du genre grammatical, où bien des détails ne sont pas encore élucidés, se laisse déterminer clairement en ses traits essentiels. Plus on l'examine, et mieux on voit que la différence entre « animé » (masculin-féminin) et « inanimé » et celle entre masculin et féminin sont hétérogènes.

COMMENT
LES MOTS CHANGENT DE SENS[1]

I

Le langage a pour première condition l'existence des sociétés humaines dont il est de son côté l'instrument indispensable et constamment employé ; sauf accident historique, les limites des diverses langues tendent à coïncider avec celles des groupes sociaux qu'on nomme des nations ; l'absence d'unité de langue est le signe d'un État récent, comme en Belgique, ou artificiellement constitué, comme en Autriche ; le langage est donc éminemment un fait social. En effet, il entre exactement dans la définition qu'a proposée Durkheim ; une langue existe indépendamment de chacun des individus qui la parlent, et, bien qu'elle n'ait aucune réalité en dehors de la somme de ces individus, elle est cependant, de par sa généralité, extérieure à chacun d'eux ; ce qui le montre, c'est qu'il ne dépend d'aucun d'entre eux de la changer et que toute déviation individuelle de l'usage provoque une réaction ; cette réaction n'a le plus souvent d'autre sanction que le ridicule auquel elle expose l'homme qui ne parle pas comme tout le monde ; mais, dans les États civilisés modernes, elle va jusqu'à exclure des emplois publics, par des examens, ceux qui ne savent pas se conformer au bon usage admis par un groupe social donné. Les caractères d'extériorité à l'individu et de coercition par lesquels Durkheim définit le fait social apparaissent donc dans le langage avec la dernière évidence.

1. Extrait de l'*Année sociologique* (1905-1906).

Néanmoins la linguistique est demeurée jusqu'à présent à l'écart de l'ensemble des études sociologiques qui se constituent si activement, et, ce qui est plus grave, étrangère presque à toute considération systématique du milieu social où se développent les langues. Cet état de choses, au premier abord surprenant et paradoxal, s'explique quand on se rend compte de la manière dont s'est créée la linguistique; les langues ne sont pas étudiées d'ordinaire pour elles-mêmes; toutes les fois qu'on les a étudiées, ç'a été en vue de la récitation correcte d'un rituel religieux, ou de l'intelligence de vieux textes religieux ou juridiques, ou pour entendre des langues étrangères, ou enfin pour parler ou écrire correctement la langue d'un grand groupe social, devenue différente de la langue de tous les jours et surtout de celle des diverses parties du groupe; on n'étudie que les langues qu'on ne parle pas naturellement, et pour arriver à les pratiquer. L'objet premier de l'étude linguistique a été partout une pratique, et l'on a été conduit ainsi à envisager, non pas les procès par lesquels se maintiennent et développent les langues, mais des faits concrets : la prononciation, les mots, les formes grammaticales et les agencements de phrases.

La linguistique a gagné à cela d'être une étude strictement objective et méthodiquement poursuivie, en un temps où la plupart des autres sciences sociales n'existaient pas ou bien n'étaient encore que de vagues idéologies ; mais tout ce que l'on peut obtenir si l'on ne sort pas de cette considération étroite des faits de langue, c'est de constater des rapports plus ou moins définis de simultanéité ou de succession entre ces faits sans jamais arriver à déterminer quelles sont les conditions générales qui en règlent l'apparition et le devenir, c'est-à-dire sans en jamais déterminer les causes.

Un grand pas a été fait, et la linguistique a déjà échappé aux limites de l'ancienne grammaire, quand on s'est attaché à définir, d'une part, les conditions anatomiques et physiologiques de l'articulation, de l'autre, les phénomènes psychiques qui interviennent dans le langage humain. Par là on parvient à se rendre compte de la raison d'être d'un grand nombre de faits linguistiques qui

relèvent immédiatement ou de la physiologie ou de la psychologie. Mais dès l'abord il apparaît qu'on ne saurait expliquer les faits uniquement à l'aide de considérations physiologiques et psychologiques ; les procédés par lesquels se réalisent les faits de langue sont devenus en partie plus clairs, mais les causes qui les déterminent sont toujours également obscures ; on voit mieux comment les langues se développent ; mais on continue d'ignorer quelles actions déterminent les innovations et les conservations dont l'ensemble constitue l'histoire du langage. Et il n'y a là rien que de naturel ; si le milieu dans lequel évolue le langage est un milieu social, si l'objet du langage est de permettre les relations sociales, si le langage n'est maintenu et conservé que par ces relations, si enfin les limites des langues tendent à coïncider avec celles des groupes sociaux, il est évident que les causes dont dépendent les faits linguistiques doivent être de nature sociale, et que seule, la considération des faits sociaux permettra de substituer en linguistique à l'examen des faits bruts la détermination des procès, c'est-à-dire à l'examen des choses, l'examen des actions, à la pure constatation de rapports entre phénomènes complexes, l'analyse de faits relativement simples considérés chacun dans leur développement particulier.

Une fois le problème ainsi posé, on s'aperçoit immédiatement que des faits qui semblent identiques tant que l'on se place au point de vue purement linguistique, sont en réalité hétérogènes. Par exemple le passage du groupe français *wé* (écrit *oi*, suivant une graphie ancienne qui avait déjà cessé d'être exacte au XIII^e^ siècle) à *wa* dans des cas tels que *moi*, *roi*, *boire*, etc., est à Paris le résultat d'un procès phonétique spontané, et qui a dû se réaliser d'une manière indépendante et nécessaire en chacun des sujets qui y ont appris à parler à une certaine date ; ailleurs cette même substitution s'est réalisée par imitation de la langue parisienne et est un fait d'emprunt ; elle peut alors parvenir à la même extension qu'à Paris ; mais le phénomène est d'ordre différent ; le linguiste pur sera sujet à confondre les deux types de faits, et il est même inévitable qu'il lui arrive de les confondre là où il n'est pas renseigné sur la façon dont le résultat iden-

tique a été obtenu sur les deux domaines considérés ; mais, s'il essaie de déterminer les causes, il ne le pourra qu'en séparant rigoureusement les deux procès, et que là où il a le moyen de les isoler ; car, d'un côté, il est en présence du type des innovations phonétiques spontanées dont le mode physiologique de développement est déjà connu avec une grande précision dans un grand nombre de cas et dont on peut même déterminer les modalités générales comme l'a fait pour certaines transformations M. Grammont, mais dont les causes efficientes n'en sont pour cela ni moins obscures ni moins énigmatiques ; et de l'autre côté, il est en présence de la substitution du français aux parlers locaux, fait historique dont les raisons immédiates sont claires et qui rentre dans le type général de la substitution des grandes langues communes de civilisation aux langues particulières de petits groupements locaux.

Le second phénomène manifeste la tendance qui entraîne les membres d'un même ensemble social à se conformer les uns aux autres en tout ce qui est utile à l'exercice de leurs fonctions communes. Le premier phénomène seul, la transformation spontanée, est à expliquer par une action identique qui a dû s'exercer d'une même manière chez tous les enfants nés à Paris durant un certain laps de temps. La distinction des deux procès est essentielle ici ; il est clair en effet qu'on n'a chance d'arriver à déterminer la nature de cette action directe que si l'on a d'abord réussi à la localiser d'une manière précise.

II

Le groupe de faits linguistiques où l'action de causes sociales est dès maintenant reconnue de la manière la plus certaine et le plus exactement déterminée est celui des innovations apportées au sens des mots[1]. Mais, conformément au principe qui vient d'être posé, celui de la distinction des procès, il convient de ne

1. On trouvera l'essentiel de la bibliographie et un bref historique de la sémantique dans un article de M. Jaberg, *Zeitschrift für romanische Philologie*, vol. XXV, p. 561 et suiv.

pas envisager tous les changements de sens d'une manière globale.

La première classification des changements de sens a été naturellement une classification logique ; on s'est demandé comment avaient varié l'extension et la compréhension des mots ; et l'on a présenté les changements de sens comme s'ils étaient l'effet des diverses sortes de métaphores. Le petit livre d'Arsène Darmesteter sur la *Vie des mots* est encore tout dominé par ces conceptions *a priori*.

Mais Michel Bréal, dans un compte rendu, a fait dès l'abord remarquer ce qu'il y a de scolastique dans ce procédé et a mis en évidence les réalités psychiques et sociales qui se cachent sous ces abstractions (voir l'article sur l'*Histoire des mots*, reproduit dans l'*Essai de sémantique*, 3e édition, p. 279 et suiv.). Depuis, ces observations ont été reprises par Bréal dans son *Essai de sémantique* et développées avec la finesse et le sens de la réalité qui caractérisaient l'auteur, mais sans recherche d'un système complet et fermé.

D'autre part, et plus récemment, Wundt, dans sa *Sprache*, consacrait aux changements de sens un long chapitre et montrait par quel jeu complexe d'associations et d'aperceptions les mots changent de sens, substituant d'une manière définitive aux subdivisions *a priori* des logiciens l'examen détaillé de la réalité psychique, et rendant impossible de parler désormais des métaphores du langage d'une manière vague, comme on le fait encore trop souvent. Mais Wundt, lui-même, ne conteste pas que l'association est loin de tout expliquer, et il serait aisé de montrer que, si elle est toujours l'élément fondamental des faits psychiques qui interviennent dans les changements de sens, elle n'est nulle part la cause efficiente qui les détermine ; ce qui fait que les études sur le développement du sens des mots, malgré de nombreuses tentatives, n'ont pas encore abouti à une théorie complète, c'est qu'on a voulu deviner les faits et qu'on ne s'astreint pas à suivre l'histoire des mots, et à tirer de l'examen de cette histoire des principes fixes ; or, nulle part moins qu'en sémantique, on ne peut déterminer *a priori* les

conditions de production des phénomènes ; car en aucune partie de la linguistique les conditions ne sont plus complexes, plus multiples et plus variées suivant les cas.

Néanmoins, il est permis de dire que, si, faute de renseignements suffisants, il est souvent — et peut-être même le plus souvent — impossible de déterminer les conditions d'un changement de sens particulier, les causes générales de ces changements sont maintenant connues pour l'ensemble, et il suffit de classer systématiquement les faits observés et les explications certaines qu'on en a trouvées, pour reconnaître que, sous le nom de changements de sens, on réunit des faits de natures bien distinctes, relevant de procès entièrement différents les uns des autres, et dont l'étude ne saurait par suite former un chapitre unique de la linguistique.

Avant d'énumérer les procès qui aboutissent aux changements de sens, il importe d'ailleurs de rappeler que les phénomènes linguistiques ont une spécificité caractéristique et que les causes efficientes qui vont être examinées n'agissent pas seules, qu'elles interviennent seulement au milieu de groupes de faits d'une nature spéciale qui sont les faits linguistiques.

Il faut tenir compte tout d'abord du caractère essentiellement discontinu de la transmission du langage : l'enfant qui apprend à parler ne reçoit pas la langue toute faite : il doit la recréer tout entière à son usage d'après ce qu'il entend autour de lui, et c'est un fait d'expérience courante que les petits enfants commencent par donner aux mots des sens très différents de ceux qu'ont ces mêmes mots chez les adultes dont ils les ont appris [1]. Dès lors, si l'une des causes qui vont être envisagées vient à agir d'une manière permanente, et si, par suite, un mot est souvent employé d'une manière particulière dans la langue des adultes, c'est ce sens usuel qui s'impose à l'attention de l'enfant, et le vieux sens du mot, lequel domine encore dans l'esprit des

1. Sur ce sujet, on peut renvoyer notamment au très intéressant chapitre du livre de M. Pavlovitch, *Le langage enfantin*, en particulier, p. 110 et suiv. et 116 et suiv.

adultes, s'efface dans la génération nouvelle ; soit, par exemple, le mot *saoul* dont le sens ancien est « rassasié » ; on en est venu à appliquer ce mot aux gens ivres, qui sont « rassasiés de boisson » ; les premiers qui ont ainsi employé le mot *saoul* s'exprimaient avec une sorte d'indulgence ironique et évitaient la brutalité du nom propre *ivre*, mais l'enfant qui les entendait associait simplement l'idée de l'homme ivre à celle du mot *saoul*, et c'est ainsi que *saoul* est devenu le synonyme du mot *ivre* qu'il a même remplacé dans l'usage familier ; par là même le mot *saoul* est celui qui maintenant exprime la chose avec le plus de crudité. Cette discontinuité de la transmission du langage ne suffirait à elle seule à rien expliquer, mais, sans elle, toutes les causes de changement auraient sans doute été impuissantes à transformer le sens des mots aussi radicalement qu'il l'a été dans un grand nombre de cas : d'une manière générale d'ailleurs, la discontinuité de la transmission est la condition première qui détermine la possibilité et les modalités de tous les changements linguistiques ; un théoricien est même allé jusqu'à vouloir expliquer par la discontinuité tous les changements linguistiques (voy. E. Herzog, *Streitfragen der romanischen Philologie*, I).

En ce qui concerne spécialement le changement de sens, une circonstance importante est que le mot, soit prononcé, soit entendu, n'éveille presque jamais l'image de l'objet ou de l'acte dont il est le signe ; comme l'a si justement dit M. Paulhan cité par M. Leroy, *Le langage*, p. 97 : « comprendre un mot, une phrase, ce n'est pas avoir l'image des objets réels que représente ce mot ou cette phrase, mais bien sentir en soi un faible réveil des tendances de toute nature qu'éveillerait la perception des objets représentés par le mot ». Une image aussi peu évoquée, et aussi peu précisément, est par là même sujette à se modifier sans grande résistance.

Tous les changements de forme ou d'emploi que subissent les mots contribuent indirectement aux changements du sens. Aussi longtemps qu'un mot reste associé à un groupe défini de formations, il est tenu par la valeur générale du type, et sa signification garde par suite une certaine fixité ; mais, si pour

quelque raison que ce soit, le groupe se disloque, les divers éléments qui le constituent, n'étant plus soutenus les uns par les autres, sont exposés à subir l'action des influences diverses qui tendent à modifier le sens.

Soit par exemple l'adjectif latin *uiuus* : il est en latin inséparable du verbe *uiuere* « vivre », du substantif *uita* « vie », etc., et, par suite, ne saurait d'aucune manière perdre le sens de « vivant ». Mais du jour où la prononciation a, comme en français, séparé l'adjectif *vif* du verbe *vivre* et où la communauté radicale avec le mot *vie* a cessé d'être perceptible, une nuance de sens qui existait déjà en latin, celle de « mobile, animé », a pu prendre le dessus.

Un mot tel que *tegmen*, qui relève d'un type de formation clair et productif en latin, est par là même inséparable du verbe *tegere* « couvrir » et garde le sens général de « couverture ». Au contraire un substantif tel que *tectum*, dont le type de formation n'est plus productif en latin, peut recevoir un sens spécialisé, celui de « toit » ; un autre substantif, appartenant à un type de formation également improductif dans la même langue, *tegula*, a pris un sens plus étroitement spécialisé encore, celui de « tuile » ; enfin *toga*, qui est une formation très ancienne et presque unique en son genre en latin, a le sens le plus éloigné de celui du groupe principal constitué par *tegere*, *tegmen*, et désigne une sorte de vêtement.

En latin, le mot *captiuus* « prisonnier » était étroitement associé à *capere* « prendre », *captus* « pris », etc., et le sens de « captif » ne pouvait par suite être perdu de vue ; mais *capere* a en partie disparu, en partie subsisté avec des significations particulières, et ce sont les représentants de *prehendere* qui expriment l'idée de « prendre » dans les langues romanes ; dès lors *captiuus* était à la merci des actions extérieures, et le mot prend le sens de « misérable, mauvais » dans l'italien *cattivo*, le français *chétif* (provincial *cheti*, signifiant « mauvais » dans une grande partie de la France).

En allemand le mot *schlecht* dont le sens était « uni, simple » a eu, sous l'influence de *schlichten* « unir, aplanir, débrouiller »,

un doublet *schlicht* ; *schlicht* étant associé à *schlichten*, a gardé le sens ancien ; mais *schlecht*, devenu un adjectif isolé, a subi un fort changement ; *ein schlechter mann* « un simple homme du commun », par opposition aux gens qui occupent un rang plus ou moins élevé ; dans une société aristocratique comme celle du XVIII^e^ siècle, où les rangs étaient bien marqués, celui qui était un *schlechter mann* était peu considéré, c'était un homme de peu, un homme sans valeur, et le mot *schlecht* a ainsi suivi la voie qu'avait suivie *captivus* en roman ; il a fini par signifier « mauvais » tout simplement, et ce sens est entièrement fixé dès le début du XIX^e^ siècle.

Le mot français dialectal *maraud* « matou » a fourni un verbe *marauder* « faire le matou » ; en Berry, où le mot *maraud* tend à disparaître, le verbe dérivé *marauder* qui signifiait d'abord « miauler bruyamment », a été appliqué à l'acte de « pleurer avec bruit et d'une manière désagréable » (employé avec une intention plutôt méprisante) ; le français littéraire, où *maraud* n'a jamais existé, a emprunté *marauder* au sens de « voler » avec une nuance particulière ; ni l'un ni l'autre de ces développements de sens n'aurait sans doute abouti aussi complètement dans des parlers où *maraud* « matou » aurait existé (sur les faits voir Sainean, *La création métaphorique en français et en roman*, I, [Halle, 1905], p. 73 et 84).

Les exemples de ce genre sont innombrables.

Mais qu'il s'agisse de la discontinuité de la transmission du langage ou de l'isolement de certains mots, les conditions linguistiques considérées ne sont jamais que des conditions en quelque sorte négatives ; elles créent la possibilité linguistique du changement de sens, mais elles ne suffisent pas à le déterminer ; elles sont des conditions nécessaires, mais non pas des conditions suffisantes, et il reste à mettre en évidence les causes efficientes des innovations.

Les causes générales qui peuvent servir à expliquer les changements de sens semblent pouvoir être ramenées à trois grands types irréductibles les uns aux autres, et qui constituent trois sortes d'actions différentes ; le résultat est, dans les trois cas, un

changement de sens, et pour cette raison, le linguiste est disposé à les grouper ; mais les trois procès sont spécifiquement distincts et n'ont en réalité rien de commun que le résultat, si bien que, dans une étude réellement scientifique, il y a lieu de les traiter séparément.

Quelques changements, en nombre assez restreint du reste, procèdent de conditions proprement linguistiques : ils proviennent de la structure de certaines phrases, où tel mot paraît jouer un rôle spécial. Ainsi, dans les phrases négatives, interrogatives ou conditionnelles, un mot vague comme *homme, chose* se trouve souvent avoir une valeur tout à fait indéfinie ; ainsi qu'on l'a déjà noté, les mots n'éveillent en général pas expressément l'image des objets auxquels ils sont associés ; et dans des tours de ce genre, très vagues par eux-mêmes et rendus plus inexpressifs encore par la fréquente répétition, aucune image n'est évoquée, ni chez celui qui parle, ni chez celui qui écoute ; le mot arménien moderne *marth* « l'homme », dans une phrase telle que *marth tch ga* « nul homme n'est ici (il n'y a personne) » ou *marth egaw* « un homme est venu ? » (quelqu'un est-il venu ?), a déjà la valeur d'un indéfini pur et simple ; le mot *manna* « homme » s'emploie de la même manière dans les textes gotiques, les plus anciens textes germaniques suivis qu'on possède ; le mot « homme » est susceptible d'acquérir ainsi la valeur d'un indéfini, et c'est par ce procédé que le français *on* (continuation du latin *homo*), l'allemand et l'anglais *man* (correspondant au gotique *manna*) ont pris leur sens caractéristique. Le mot latin *alter* signifiait « autre », quand il s'agit de deux objets, par conséquent « second, l'un des deux » ; mais dans une phrase négative, *alter* ne se distingue pas essentiellement pour le sens de *alius* « autre par rapport à plus de deux » ; la phrase d'Ovide : *neque enim spes altera restat* peut se traduire à volonté : « il n'y a pas un second espoir », ou « il n'y a pas d'autre espoir », sans que le sens soit au fond changé ; le mot *alter* a pris dans ce type de phrases la valeur de *alius* ; cette valeur a été transportée dans des phrases quelconques, et

les langues romanes, laissant tomber *alius*, n'ont conservé que *alter*, pour exprimer le sens de « autre » ; la disparition du comparatif et du superlatif avait du reste fait perdre l'habitude d'opposer la comparaison entre deux objets (type *ualidior manuum* « la plus forte des deux mains ») à la comparaison entre plusieurs (*ualidissimus uirorum* « le plus fort des hommes »). De même, sous l'influence de *ne*, les mots français *pas*, *rien*, *personne* ont pris, dans les phrases négatives, une valeur négative, si bien que la négation *ne* est devenue inutile dans le français actuel et que *pas*, *rien*, *personne* sont négatifs par eux-mêmes dans la langue familière et courante. Le mot latin *magis* « plus, de plus, bien plus », placé en tête de la phrase, comme il arrive déjà en latin, fait l'effet d'une liaison entre deux phrases et devient le français *mais*. On le voit, tous ces procès purement linguistiques aboutissent moins à créer un changement de sens qu'à transformer des mots à sens concret en simples outils grammaticaux, en éléments de construction de la phrase. C'est une conséquence immédiate de la nature même du procès en question.

Inversement, les catégories grammaticales servent parfois à transformer le sens d'un mot : le latin *homo* servait à indiquer l' « homme » en tant qu'être humain, sans acception de sexe ; mais le genre grammatical de *homo* était le masculin qui, là où il a une valeur définie, a celle de désigner le sexe mâle ; le représentant roman de *homo* a été amené ainsi à joindre au sens de « être humain » celui de « homme de sexe masculin », et le mot *uir*, qui avait ce sens en latin ancien, a été éliminé. — Une même racine fournit en grec un aoriste signifiant « voir », ἰδεῖν, et un parfait signifiant « je sais », οἶδα ; ces deux sens sont anciens, car ils se retrouvent l'un dans latin *uidere*, etc., l'autre dans sanskrit *veda* « je sais », gotique *wait* (allemand *weiss*), etc., et le slave oppose également *vidĕti* « voir » à *vědĕti* « savoir » ; ils tiennent à ce que l'aoriste indiquant l'action pure et simple se prête à noter une simple sensation : « voir », tandis que le parfait, qui indique le résultat acquis d'un acte antérieur, convient pour signifier « savoir ».

Ces cas où l'agent essentiel du changement est la forme

grammaticale sont d'une espèce assez rare, car les catégories grammaticales qui répondent à quelque réalité objective sont en petit nombre, et par suite les conditions de réalisation de ces procès ne se rencontrent pas très souvent ; mais la forme grammaticale du mot est partout l'un des éléments dont dépend le changement ou le maintien du sens.

Un second type de changements de sens est celui où les choses exprimées par les mots viennent à changer. Les mots français *père* et *mère* sont la continuation exacte des mots indo-européens qui indiquaient le père et la mère, et pourtant les mots français ne sont pas associés aux mêmes représentations que les mots indo-européens correspondants ; ces mots indo-européens désignaient des relations sociales définies bien plutôt qu'ils ne visaient la relation de paternité et de maternité physiologiques, cette dernière étant exprimée par des mots qui sont en latin *genitor* et *genetrix* ; mais la structure sociale ayant changé, la famille patriarcale indo-européenne ayant disparu, les mots *père* et *mère* expriment avant tout la paternité et la maternité physiques ; et dès lors on est conduit à appliquer les mots de *père* et *mère* aux animaux ; en français populaire, un *père* est un « mâle », et une *mère* une « femelle », et ce sens est si complètement développé qu'il y a des patois français où les formes locales de *père* et *mère* signifient simplement « mâle » et « femelle » d'animaux, et où l'on recourt aux formes françaises communes — qui semblent plus élégantes et conformes à la dignité des parents — pour désigner proprement le « père » et la « mère » ; dans les anciennes langues indo-européennes, les mots qui correspondent au latin *pater* et *mater* n'admettent pas cet emploi ; ils indiquent une situation sociale, une dignité, et ils ont une valeur religieuse, qui apparaît clairement dans le t. *Juppiter* (ancien *dyeu-pater* « Père-Cid »).

Un autre exemple montre à quels changements étranges peut entraîner la variation des choses. La peine infligée aux criminels français à un certain moment a été de les envoyer ramer sur les *galères* du roi ; la locution *envoyer aux galères* était donc une

manière d'exprimer la condamnation à une peine grave ; on a cessé ensuite d'employer les rames sur les vaisseaux, les criminels ont été envoyés dans des bagnes de terre ferme ; mais l'expression *envoyer aux galères* a subsisté, et un *galérien* a été un forçat qui subissait sa peine dans une sorte de prison ; le terme est maintenant en voie de disparition, mais dans la mesure où l'on s'en sert encore, il n'a plus d'autre valeur que celle de « forçat ».

Les changements de ce genre atteignent constamment presque tous les mots ; mais on ne les remarque que lorsqu'ils présentent quelque chose de singulier et d'étrange : on dit du *papier* (latin *papyrum*) de chiffons ; la *plume* de fer s'est substituée à la *plume* d'oie sans que le nom ait varié ; et ainsi de suite : les changements des choses ne se traduisent que d'une manière restreinte par des changements des mots : car les mots étant associés à des représentations toujours très complexes s'associent facilement à des représentations qui ont avec celles d'une génération précédente quelques traits communs. Et c'est ainsi que la variation de sens de beaucoup de mots, c'est-à-dire au fond la variation des notions auxquelles est associé le nom donné, traduit des changements sociaux plus profonds : toute l'histoire de la société grecque se réfléchit indirectement dans le contraste entre le compagnon de guerre et d'expédition maritime de l'époque homérique, l'ἑταῖρος, et la courtisane athénienne ou alexandrine, l'ἑταίρα.

De cette catégorie il faut rapprocher les changements qui ont lieu quand un mot en doit remplacer un autre frappé de quelque « tabou », ou, ce qui est un fait d'espèce voisine, éliminé pour quelque raison de convenance : si les noms propres de la prostituée sont évités par convenance, on est conduit à associer à la prostituée le nom de la femme mariée ; et c'est ainsi que *garce*, puis *fille* ont été appelés successivement à fournir le nom de la fille publique ; il y a ici une application d'un nom à un objet dont il n'était pas le nom propre, mais qui a été attribué à cet objet par un acte exactement comparable à celui qui a fait nommer *plume* la pointe de fer substituée à la plume d'oie taillée qu'on employait antérieurement ; la cause initiale est ici de

nature sociale, mais cette cause sociale agit à peu près de la même manière qu'agit le changement de la réalité désignée par le nom.

Un même mot change de sens suivant les lieux; ainsi un mot indo-européen **prtu-*, qui désigne un « endroit par où on peut passer », signifie, suivant le cas, un pont, une porte, un gué (ces trois sens sont attestés en ancien iranien, dans la langue de l'Avesta); c'est le hasard des circonstances locales qui fait que le latin ne garde *portus* qu'au sens de « port » (tandis que le mot voisin *porta* prend celui de « porte »), et que le gaulois *ritu-* dans *Ritu-magus* « champ du gué », le vieux gallois *rit* et l'anglo-saxon *ford*, le vieux haut allemand *furt* (qui sont le même mot) conservent seulement la valeur de « gué ».

Les développements de sens reflètent l'organisation sociale, l'organisation domestique. Il est intéressant par exemple de voir comment le mot qui signifie « dehors » provient du nom de la porte, ainsi en latin *foras* et *foris*, en grec : θύραζε, θύρασι, θύρηφι, en arménien : *durs*, en persan : *dar*; et ceci coïncide avec le fait que simultanément « dehors », se dit « aux champs », c'est-à-dire « hors de la maison » dans irlandais *immag* « foras » et *immaig* « foris », à côté de *mag* « champ », dans breton *erméaz*, gallois *i maes* à côté de *méaz*, *maes* « champ », dans lituanien *laukan*, *lauke*, à côté de *laukas* « champ » et dans arménien *artakhs* à côté de *art* « champ »; ce sont les expressions qui avaient cours dans chacune des grandes familles qui étaient l'unité sociale par excellence; on y opposait l'enclos familial, le *dvor* slave, à tout ce qui était en dehors, notamment aux champs. — Un mot tel que le latin *sponsa* « promise » prend le sens de « fiancée », d'où dans certaines langues romanes celui d'« épouse », parce que le verbe latin *spondeo* « je promets » est le terme rituel prononcé par le père pour répondre « oui » à un prétendant à la main de sa fille.

Et ceci amène naturellement à envisager l'ordre des causes qui forme l'objet principal de la présente étude, la répartition des hommes de même langue en groupes distincts : c'est de cette hétérogénéité des hommes de même langue que procèdent le

plus grand nombre des changements de sens, et sans doute tous ceux qui ne s'expliquent pas par les causes précitées.

III

L'action de la division des hommes en classes distinctes sur le sens des mots a déjà été souvent signalée par les auteurs qui ont écrit sur la sémantique ; et Bréal en particulier l'a exprimée avec une grande précision : « A mesure qu'une civilisation gagne en variété et en richesse, les occupations, les actes, les intérêts dont se compose la vie de la société se partagent entre différents groupes d'hommes ; ni l'état d'esprit, ni la direction de l'activité ne sont les mêmes chez le prêtre, le soldat, l'homme politique, l'agriculteur. Bien qu'ils aient hérité de la même langue, les mots se colorent chez eux d'une nuance distincte, laquelle s'y fixe et finit par y adhérer... Au mot d'*opération*, s'il est prononcé par un chirurgien, nous voyons un patient, une plaie, des instruments pour couper et tailler ; supposez un militaire qui parle, nous pensons à des armées en campagne ; que ce soit un financier, nous comprenons qu'il s'agit de capitaux en mouvement ; un maître de calcul, il est question d'additions et de soustractions[1]. Chaque science, chaque art, chaque métier, en composant sa terminologie marque de son empreinte les mots de la langue commune » (*Essai de sémantique*, 3e édit., p. 285 et suiv. ; voir surtout les chapitres de la *Polysémie*, p. 143 et suiv., et *D'un cas particulier de polysémie*, p. 151 et suiv.). On trouvera des observations analogues de L. Duvau, *Mémoires de la Société de linguistique*, XIII, 234 et suiv., de M. Meringer, *Indogermanische Forschungen*, XVII, de M. Schuchardt, dans son travail sur *trouver*, *Sitzungsberichte* de l'Académie de Vienne, phil. hist. cl., vol. CXLI (année 1899) ; on consultera aussi Roques, *Méthodes étymologiques*, *Journal des savants*, août 1905).

1. Bréal aurait pu ajouter, entre autres exemples, qu'aux entrepôts de vin de Bercy, *opération* désigne un mélange de vins, et que tout le monde y entend par *vin d'opérations* un vin qui sert à des coupages.

Dans sa publication dédiée à Adolf Mussafia (Graz, 1905), M. Schuchardt écrit : « Bien que l'origine de tous (les mots signifiant « devoir » [all. *müssen*] en italien dialectal) ne soit pas éclaircie, il semble qu'il s'y réfléchisse surtout des différences sociales. Le devoir de l'esclave n'est pas celui du maître, l'esclave aura aussi facilement un *mihi ministerium est* aux lèvres que le maître un *mihi calet.* »

Le fait fondamental est donc qu'un mot qui, dans la langue commune d'une société, a un sens étendu s'applique, dans un des groupes restreints qui existent à l'intérieur de cette société, à des objets plus étroitement déterminés, et inversement ; M. Meringer dit très bien, dans *Indogermanische Forschungen*, XVIII, 232 : « un mot élargit sa signification quand il passe d'un cercle étroit à un cercle plus étendu ; il la rétrécit quand il passe d'un cercle étendu à un cercle plus étroit ». L'exemple du mot *opération* définit assez le principe pour qu'il soit inutile d'en ajouter d'autres ; aussi bien le fait est-il d'expérience courante. Chaque groupe d'hommes utilise d'une manière particulière les ressources générales de la langue.

Les groupes où le sens des mots se précise ainsi ne sont pas seulement des groupes professionnels ; tout ensemble d'individus, qui a, à quelque point de vue que ce soit, des relations spéciales à l'intérieur d'une société a, par là même, des notions spéciales et obéit à des convenances spéciales au petit groupe qu'il constitue, soit transitoirement, soit d'une manière permanente ;. or, la signification d'un mot est définie par l'ensemble des notions auxquelles est associé le mot, et les associations diffèrent évidemment suivant le groupe où le mot est employé. Le vocabulaire des femmes n'est pas identique à celui des hommes : le mot *habiller* a, en français, une tout autre valeur chez les femmes que chez les hommes, parce qu'il s'applique à un acte dont le caractère et l'importance sont entièrement différents. Ailleurs c'est par convenance que les femmes s'expriment autrement que les hommes : il y a par exemple un dialecte serbe où les femmes évitent le nom propre du bœuf, *kurjak*, employé par les hommes, parce que ce mot a en même temps le sens de

« *penis* », et recourent à d'autres mots. On emploie partiellement une terminologie spéciale à la caserne, dans un groupe d'étudiants, dans un groupe sportif ; et, il importe de le noter, les mêmes individus appartiennent simultanément ou successivement à plusieurs des groupes en question, si bien qu'ils subissent à la fois ou à divers moments de leur vie des influences diverses.

Les hommes qui exercent une même profession ont à désigner un grand nombre d'objets et de notions pour lesquels la langue commune n'a pas de noms parce que le commun des hommes ne s'en occupe pas. Beaucoup de ces désignations sont obtenues en attribuant à des objets le nom d'autres objets avec lesquels ceux-ci ont une ressemblance plus ou moins lointaine ; on désigne ainsi sous le nom de *chèvre* telle machine servant à porter ; en anglais, *cat* « chat » est aussi un crampon qui sert à saisir l'ancre (d'après les griffes du chat, etc.). On n'entend marquer par là que des analogies vagues, et très souvent, au lieu de recourir au mot lui-même, on se sert d'un dérivé : le *chevalet* est autre chose que le *cheval*, la *manette* autre chose que la *main* ; ce procédé de dérivation est de règle en russe, où le « bec » d'une cafetière est un *nosik* et non un *nos* « nez » (voir Boyer et Spéranski, *Manuel de russe*, p. 113, n. 4).

Quelle que soit la nature du groupe considéré, le sens des mots est sujet à y varier non seulement en raison des circonstances spéciales qui le déterminent, comme il arrive par exemple pour le mot *opération*, mais aussi en raison de ce qu'il s'agit d'un groupe plus ou moins isolé du reste de la société, plus ou moins fermé, plus ou moins autonome ; car la variation du vocabulaire ne se limite pas à ce qu'exige la nature même du groupe ; elle est grossie intentionnellement par suite de la tendance qu'a chaque groupe à marquer extérieurement son indépendance et son originalité : tandis que l'action de la société générale tend à uniformiser la langue, l'action des groupements particuliers tend à différencier, sinon la prononciation et la grammaire, qui restent sensiblement unes, du moins le vocabulaire, des individus qui y prennent part. Il y a là deux tendances antagonistes qui résultent immédiatement et du

caractère de la langue générale et du rôle spécial des langues particulières.

Les langues de groupes particuliers deviennent ainsi des « argots », et ces argots eux-mêmes se constituent parfois en langues artificielles, par des altérations systématiques, ainsi en France le jargon des bouchers, le *loucherbème* : ce qui montre bien que le fait est naturel, c'est qu'il se retrouve dans des langues tout à fait différentes. M. Chéron décrit ainsi les argots des marchands de porcs, des marchands de grain, des sampaniers, des chanteuses, etc., du Tonkin, qui sont autant de déformations de l'annamite (voir *Bulletin de l'École française d'Extrême-Orient*, V, 47 et suiv.).

L'action de la tendance aux innovations de sens dans les langues particulières est d'ailleurs facilitée par plusieurs circonstances. Dans un groupe limité, il est souvent question des mêmes choses ; les associations d'idées sont les mêmes chez les divers individus, et l'on s'y entend sans avoir besoin de préciser ; d'autre part, ce qui, pour une personne étrangère au groupe, serait obscur est clair pour les membres du groupe dès l'instant que certains procédés d'expression ont commencé d'y avoir cours, qu'une manière s'y est créée.

Ainsi s'explique le trait caractéristique des changements de sens argotiques que M. Schwob et G. Guieysse ont mis en évidence (*Mémoires de la Société de linguistique*, VII, 33 et suiv.) : la dérivation synonymique [1]. Si un mot A a simultanément deux significations, l'une *x* dans la langue générale, l'autre *y* en argot, tous les synonymes approximatifs du mot A de la langue générale au sens *x* seront admis en argot à avoir la signification *y* de l'argot ; par exemple, si *polir* est employé en argot au sens de « voler », qui se rencontre déjà chez Villon, on pourra employer de même *fourbir, brunir, sorniller, nettoyer* ; si *battre* signifie une fois « tromper » en argot, on pourra donner le même sens à *taper, estamper*, etc. Le souci de demeurer inintelligible au vul-

1. Dans *Le poilu tel qu'il se parle* (Paris, 1919), M. G. Esnault a donné beaucoup d'exemples, bien analysés, de ce procédé, et d'autres faits de sémantique, d'origine sociale.

gaire est pour beaucoup dans le développement considérable qu'a reçu le procédé en argot ; mais le principe même n'est pas propre à l'argot entendu au sens étroit, et le procédé se retrouve, en une mesure plus ou moins étendue, dans toutes les langues de groupes particuliers ; par exemple, dans un groupe où l'on s'est mis à employer des adverbes tels que *terriblement* pour exprimer ce que la langue commune indique par *très*, on est conduit à employer à l'occasion tous les synonymes approximatifs tels que *effroyablement*, *redoutablement*, ou des adverbes de même sorte ; c'est sans doute à quelque synonymie de ce genre que les formes négatives, *pas*, *point*, *mie* du français, doivent leur origine ; du jour où l'un de ces mots a pris une valeur proprement négative, il a éliminé les autres ; *mie* est sorti de tout l'usage français, *point* de l'usage dans la langue parlée, et il n'est resté que *pas*, lequel a cessé d'être une détermination de la négation pour devenir par lui-même la négation usuelle en français parlé. De pareilles modifications du sens des mots par synonymie ne s'expliquent que dans des groupes fermés ; la résistance à l'innovation linguistique, qui est chose normale dans l'ensemble du groupe social, est anéantie sur un point particulier dans le petit groupe en question où, en se singularisant à l'égard de l'ensemble, l'individu ne fait que mieux marquer sa solidarité avec le groupe étroit dont il fait partie.

L'une des causes qui font que les groupes particuliers sont éminemment propres à modifier leur vocabulaire, c'est que les éléments qui constituent chaque groupe ne sont souvent pas homogènes au point de vue linguistique et que, de plus, ils sont soumis à des influences étrangères. En effet les groupements qui se forment à l'intérieur d'une société, et notamment les groupements professionnels, sont composés de gens qui ne sont pas nécessairement issus d'une même localité, ni même d'une même région, et dont, par suite, la langue n'est pas identique : par elle-même, et sans qu'on fasse intervenir l'action de l'une quelconque des langues locales en question, cette absence d'homogénéité est évidemment une cause d'instabilité et d'incertitude, et — on ne l'a pas assez remarqué — c'est l'une des principales

causes, la principale peut-être, de tous les changements linguistiques, de ceux de la prononciation et de la grammaire comme de ceux du vocabulaire, des changements spontanés comme des emprunts.

De plus, les éléments étrangers tendent à introduire dans la langue du groupe des formes de leur propre langue : c'est ainsi que la langue des étudiants allemands renferme des mots d'origines dialectales très diverses ; dans sa *Studentensprache*, 65, M. Kluge en fournit des exemples, notamment le bas allemand *gnote* au lieu du haut allemand *genosse* « compagnon ». M. Horn fait la même remarque pour la langue des soldats allemands, dans sa *Soldatensprache*, 9 et suiv. Cette influence d'éléments étrangers se manifeste souvent par des traductions ; ainsi, dans la langue spéciale des premiers chrétiens, l' « ancien » qui était le « prêtre » était désigné par le mot πρεσβύτερος en grec ; dans le groupe des chrétiens de langue latine, où se trouvaient mêlés de nombreux éléments helléniques ou hellénisés, le mot a été gardé tel quel ; on a dit *presbiter*, qui a subsisté en français sous les formes *prêtre* et aussi *prouvoire* en vieux français.

On conçoit enfin qu'on puisse recourir à un autre procédé et que, au lieu d'employer le mot étranger, on le traduise, c'est ce qui a été fait en Orient : en Arménie le mot *erêç* « ancien » a aussi reçu le sens de « prêtre », et le géorgien fait exactement de même avec son mot signifiant « ancien ». Ou bien encore on peut, ce qui revient à peu près au même, charger un mot national d'un sens étranger ; ainsi l'anglo-saxon *eorl* « homme libre (qui va à la guerre), noble » a reçu sous Knut le sens du mot norrois *jarl* « vice-roi, gouverneur de province » qui était reconnu par le sujet parlant pour identique au mot anglais ; sous la domination normande, le même *eorl* a servi d'équivalent au français *comte*, et c'est cette valeur que *earl* a encore en anglais moderne.

De ce double procédé d'emprunt et de traduction de termes étrangers il résulte que les vocabulaires des groupes particuliers qui sont en relation avec des groupes pareils dans des pays parlant d'autres langues présentent d'ordinaire un grand nombre

de ressemblances. Le vocabulaire militaire par exemple est à peu près le même d'un bout de l'Europe à l'autre.

Le fait est particulièrement sensible dans les groupes composés de savants, ou bien où l'élément scientifique tient une place importante. Les savants, opérant sur des idées qui ne sauraient recevoir une existence sensible que par le langage, sont très sujets à créer des vocabulaires spéciaux dont l'usage se répand rapidement dans les pays intéressés. Et comme la science est éminemment internationale, les termes particuliers inventés par les savants sont ou reproduits ou traduits dans des groupes qui parlent les langues communes les plus diverses. L'un des meilleurs exemples de ce fait est fourni par la scolastique dont la langue a eu un caractère éminemment européen, et à laquelle l'Europe doit la plus grande partie de ce que, dans la bigarrure de ses langues, elle a d'unité de vocabulaire et d'unité de sens des mots. Un mot comme le latin *conscientia* a pris dans la langue de l'école un sens bien défini, et les groupes savants ont employé ce mot même en français ; les nécessités de la traduction des textes étrangers et le désir d'exprimer exactement la même idée ont fait rendre la même idée par les savants germaniques au moyen de *mith-wissei* en gotique, de *gi-wizzanî* en vieux haut-allemand (allemand moderne *gewissen*). Souvent les mots techniques de ce genre sont traduits littéralement et n'ont guère de sens dans la langue où ils sont transférés ; ainsi le nom de l'homme qui a de la pitié, latin *misericors*, a été traduit littéralement en gotique *arma-hairts* (allemand *b-arm-herzig*) et a passé du germanique en slave, par exemple russe *milo-serdyj*. Ce sont là de pures transcriptions cléricales de mots latins.

Quand, comme il est arrivé assez souvent au cours de l'histoire, les éléments dominants d'une nation ont parlé une langue différente de celle des autres groupes, les parties de la nation qui approchent immédiatement la caste dominante et qui nécessairement apprennent plus ou moins la langue de cette caste se constituent un vocabulaire où figurent un grand nombre de termes étrangers au moins pour les notions qui importent à la caste. Le nom vieil anglais de l'armée *here* a été éliminé dans le

langage des gens qui entouraient l'aristocratie normande au profit des mots d'origine française *army* et *host*.

A l'intérieur d'une langue donnée, définie par une prononciation une et surtout par l'identité des formes grammaticales, il y a en réalité autant de vocabulaires particuliers qu'il y a de groupes sociaux ayant une autonomie dans la société qui parle cette langue, et tout groupe d'hommes a ses désignations spéciales, non seulement de ce qui lui est particulier, mais aussi de nombreuses choses qui lui sont communes avec les autres membres des groupes plus étendus dont ces hommes font partie ; les exemples pourraient aisément être multipliés ; les types indiqués suffisent à fixer les idées.

Les changements de sens qui viennent d'être sommairement décrits ne restent pas confinés dans les cercles où ils se produisent. Une fois sortis des groupements particuliers où ils ne figurent que d'une façon transitoire ou périodique, les individus n'échappent pas aux habitudes qu'ils y ont contractées, et même quand ils ont affaire à des personnes étrangères aux divers groupes dont ils font eux-mêmes partie, ils restent sujets à employer les mots avec le sens que ceux-ci ont pris dans un groupe. D'autre part, s'il s'agit de groupes qui ont un prestige, notamment de groupes aristocratiques ou de groupes savants, les individus qui n'y ont pas accès se plaisent à en reproduire les usages, et notamment le vocabulaire ; ainsi des mots germaniques qui désignaient anciennement le « chef », le « seigneur », à savoir *frô* et *truhtîn*, le premier n'apparaît plus en vieux haut allemand qu'en fonction de vocatif, pour interpeller, et le second sert presque uniquement à désigner « Dieu » (le seigneur céleste) ; le chef terrestre, le seigneur est désigné par un mot calqué sur le latin *senior*, le mot *hêrro* et ce mot nouveau, emprunté par l'aristocratie germanique à la nomenclature latine, a si bien remplacé dans tout l'ensemble de l'allemand les vieux mots que, dès le XI^e^ siècle, le vieux haut allemand tend à employer *hêrro* même pour Dieu et qu'aujourd'hui seul subsiste le mot *Herr* (voir Ehrismann, *Zeitschrift fur deutsche Wortforschung*, VII, p. 173 et suiv.). Cette extension est

d'ailleurs nécessaire dans beaucoup de cas ; car c'est seulement dans les vocabulaires spéciaux que nombre de notions nouvelles ont trouvé d'abord une expression propre et exacte.

Les sens particuliers qui se sont produits dans des groupements étroits ont donc de nombreuses occasions de passer à la langue commune, soit par mode, soit par nécessité ; il y a là de véritables *emprunts* à l'intérieur d'une même langue.

Il importe de définir ici ce que l'on entend en linguistique par l'emprunt.

Soit une langue considérée à deux moments successifs de son développement ; le vocabulaire de la seconde époque considérée se compose de deux parties, l'une qui continue le vocabulaire de la première ou qui a été constituée sur place dans l'intervalle à l'aide d'éléments compris dans ce vocabulaire, l'autre qui provient de langues étrangères (de même famille ou de familles différentes) ; s'il arrive que quelque mot soit créé de toutes pièces, ce n'est, semble-t-il, que d'une manière exceptionnelle, et les faits de ce genre entrent à peine en ligne de compte. Soit par exemple le latin à l'époque de la conquête de la Gaule par les Romains et le français (c'est-à-dire la langue de Paris) au commencement du xx^e^ siècle ; il y a des mots comme *père*, *chien*, *lait*, etc., qui continuent simplement des mots latins ; il y en a comme *noyade* ou *pendaison* qui ont été faits sur sol français avec des éléments d'origine latine, et il y en a d'autres qui sont entrés à des dates diverses : *prêtre* est un mot qui est entré par le groupe chrétien à l'époque impériale romaine, sous la forme *presbyter ; guerre*, un mot germanique, apporté par les invasions germaniques, et entré dans la langue par le groupe des conquérants qui ont été maîtres du pays, à la suite de ces invasions ; *camp* est un mot italien venu au xv^e^ siècle par les éléments militaires qui ont fait les campagnes d'Italie ; *siècle* est un mot pris dès avant le x^e^ siècle au latin écrit par les clercs et qui avait disparu de la langue commune ; *équiper* est un terme de la langue des marins normands ou picards ; *foot-ball* est un terme de sport venu de l'anglais il y a peu d'années ; mais, par rapport au latin

de l'époque de César, tous les mots en question sont également empruntés, car aucun n'est la continuation *ininterrompue* de mots latins de cette date, ni ne s'explique par des formes qui se soient perpétuées dans la langue sans interruption entre l'époque de César et le commencement du xx^e^ siècle. Il n'importe pas que le mot soit emprunté à une langue non indo-européenne, comme il arrive pour *orange*, ou à une langue indo-européenne autre que le latin, comme *prêtre*, *guerre*, ou au latin écrit comme *siècle*, *cause*, ou à un dialecte roman comme *camp*, *camarade*, ou même à des parlers français plus ou moins proches du parisien, comme le mot *foin*, pris à des parlers ruraux, et que sa phonétique dénonce comme n'étant pas parisien ; en aucun cas il n'y a eu continuation directe et ininterrompue du mot latin *à Paris* depuis l'époque de César jusqu'au début du xx^e^ siècle, et ceci suffit à définir l'*emprunt* pour la période considérée de l'histoire du français parisien. Donc la notion d'*emprunt* ne saurait être définie qu'à l'intérieur d'une période strictement délimitée, et pour une population strictement délimitée.

Mais d'après ce qui a été exposé ci-dessus, un mot peut porter toutes les marques phonétiques et morphologiques auxquelles on reconnaît un mot non emprunté ; il peut même avoir subsisté sans interruption dans la langue, et être néanmoins au fond un mot *emprunté*, si, pendant un temps plus ou moins long, il n'a plus fait partie de la langue commune et s'il a été employé seulement dans des groupes sociaux particuliers. Sans parler des autres causes qui ont pu intervenir, c'est sans doute pour n'avoir subsisté que dans le langage rural que des mots latins comme *ponere* « placer », *cubare* « être couché », *trahere* « tirer », *mutare* « changer » ont pris des sens tout particuliers et techniques et ont fourni au français actuel *pondre*, *couver*, *traire*, *muer* ; c'est ainsi que, dans le patois français de Charmey (canton suisse de Fribourg), le mot *χlâ* « fleur », qui répond au français *fleur*, disparaît en ce sens, mais se maintient au sens de « crème », technique dans le parler de ce pays de laitage (voir Gauchat, *L'unité phonétique*, dans *Aus romanischen Sprachen*, Festschrift-Morf, p. 191). Au sens de « couper le blé », Furetière (cité

dans le *Dictionnaire général* de Hatzfeld, Darmesteter et Thomas, sous « scier ») indique que quelques personnes disent *soyer* ou *seier* au sens de « scier » (couper) du blé ; c'est un mot rural que l'on avait transporté à Paris, mais qui aurait tout aussi bien pu prendre cette forme à Paris. A prendre le terme dans un sens strict et rigoureux conforme à l'esprit même de la définition, les mots français *pondre*, etc., sont encore des mots *empruntés*, bien qu'ils n'aient peut-être jamais cessé d'exister à Paris et qu'ils aient tous les caractères phonétiques et morphologiques de mots français. MM. Gilliéron et Mongin ont, dans leur étude de géographie linguistique, *Scier dans la Gaule romane*, posé ce principe capital que beaucoup de mots qui ne se dénoncent pas par leur forme phonétique comme des emprunts sont néanmoins empruntés à des parlers voisins et peuvent être reconnus pour tels à certains indices ; « le mot *scier*, qui continue le latin *secare* « couper », ne représente, d'après ces auteurs, qu'un emploi tout particulier et spécialement rural de *secare*, à savoir « couper le blé (avec la faucille dentelée) » ; c'est un exemple à joindre à *pondre*, *couver*, etc.

Inversement, les mots empruntés ne le sont pas en principe par la langue générale ; c'est dans les groupements particuliers qu'on emprunte des mots, et, presque pour chacun des exemples cités ci-dessus, on aperçoit aisément quel est le groupe qui a fait l'emprunt. C'est par les milieux militaires, par les commerçants et par les prêtres que le germanique a emprunté au latin la grande masse des mots qu'il s'est assimilés à date ancienne ; c'est aussi par les milieux militaires et par les prêtres que le slave a, à date ancienne, reçu des mots du latin et du germanique ; aujourd'hui c'est surtout le monde du sport qui emprunte des mots à l'anglais, c'est le monde militaire qui en emprunte à l'allemand, etc.

L'emprunt qui est, de tous les faits linguistiques intéressant le vocabulaire, le plus important sans doute a donc avant tout des causes sociales, et comme les causes sociales sont ici évidentes et que leur action y est immédiate tandis que pour d'autres faits linguistiques elle est plus obscure et moins directe, on est allé

jusqu'à proposer de qualifier les emprunts morphologiques de faits de linguistique sociale, par opposition aux développements spontanés qui seraient individuels (Wrede, *Archiv für das Studium der neueren Sprachen*, CXI, p. 33); il n'y a pas lieu de discuter ici cette proposition dont il serait du reste aisé de démontrer l'inexactitude.

La nature et la portée de l'emprunt étant ainsi définies, on peut poser en principe que la langue commune emprunte beaucoup aux langues particulières. Le vocabulaire d'une langue telle que le français se compose pour la plus grande partie de mots empruntés. Seuls font exception les termes généraux de la langue commune, et c'est pour cette raison que les linguistes font reposer sur ce petit nombre de mots leurs théories relatives à l'histoire des langues.

Si les mots sont empruntés par la langue générale seulement pour exprimer les notions auxquelles les a associés la langue particulière par laquelle ils ont passé, il n'y a rien de plus à en dire ; ils demeurent à l'état de corps plus ou moins étrangers, de termes techniques et ne sont dans la langue commune que des éléments accessoires ; c'est le cas qui a le premier attiré l'attention, mais dont l'importance est au fond le moindre.

S'ils pénètrent vraiment dans la langue commune et y sont employés couramment, les mots empruntés ne le font qu'en subissant un changement de sens. La valeur précise et rigoureuse d'un terme tient à l'étroitesse d'un milieu où dominent les mêmes intérêts et où l'on n'a pas besoin de tout exprimer ; sorti de ce milieu étroit auquel il devait sa valeur spéciale, le mot perd immédiatement de sa précision et tend à devenir de plus en plus vague. Pour un marchand des rues, *camelote* signifie la marchandise quelconque qu'il a entre les mains (et de même pour le chiffonnier) ; en entrant dans la langue commune, le mot a pris le sens vague de « marchandise de peu de valeur, mauvaise marchandise ». Soit encore, par exemple, le mot latin *caussa* (*causa*) ; dans la langue du barreau romain, il désignait « une affaire judiciaire, un procès » ; passant de là dans la

langue commune, il n'a plus signifié qu'une « affaire », et enfin une « chose » si bien qu'il a pu s'appliquer non seulement à une affaire, mais à un « objet » et que *chose* est devenu l'un des mots les plus vagues de toute la langue française. Le même mot, emprunté, sous forme savante, à la langue spéciale de la scolastique avec le sens de « cause », qui était le sens général du mot en latin, a passé aussi dans la langue commune, mais avec une valeur de plus en plus imprécise et sert à désigner non plus la cause efficiente ou la cause finale, mais tout motif d'action : *à cause?* équivaut à « pourquoi ? » dans la langue populaire. Et ceci n'empêche pas que, au barreau, il n'existe un mot *cause*, emprunté au latin par les juristes, auquel les avocats gardent son sens premier latin « d'affaire judiciaire » ; toutefois comme le terme appartient surtout aux avocats et ne désigne souvent une affaire en tant qu'elle est confiée à un avocat, un mot *cause* signifiant « affaire à plaider » a passé en français commun, et toute personne qui défend un parti s'attache à la bonne ou à la mauvaise *cause* ; de nouveau on se trouve très loin du point de départ du sens du mot.

De même que l'emploi dans une langue particulière détermine un changement de sens, l'emprunt fait par les langues générales à une langue *particulière* en détermine donc un autre dans un sens tout différent. Et ce n'est qu'une conséquence de la manière dont s'établit le sens des mots. Wundt, *Sprache*, 2e édit., vol. II, p. 484 et suiv., montre bien comment un mot ne désigne pas nécessairement une idée générale : pour chaque individu, le mot ne désigne même le plus souvent que certains objets particuliers, qui font partie de son expérience. Mais le mot sert en même temps à d'autres membres de la communauté pour lesquels il désigne d'autres objets plus ou moins semblables ; il se dépouille par là de tout ce qu'il a de particulier, pour ne garder d'autre rôle que celui d'indiquer les seuls caractères communs à tous les objets désignés par le mot dans un groupe social donné ; l'enfant qui apprend le mot *chien* est naturellement porté à ne l'appliquer qu'au chien de la maison, et c'est seulement au fur et à mesure qu'il entend ce même mot appliqué à

d'autres animaux qu'il lui ôte son caractère concret et lui attribue une valeur générale. On voit par là que la valeur générale des mots est, dans une très large mesure, un fait social, et que la généralité du sens d'un mot a souvent chance d'être proportionnée à l'étendue du groupe : dans le patois d'un village de pasteurs, le *chien* est par excellence le chien de berger ; mais dans une langue telle que le français, le mot *chien* exclut toute association spéciale à un type déterminé et désigne d'une manière abstraite une espèce animale dont les variétés sont nombreuses et diverses.

Il apparaît ainsi que le principe essentiel du changement de sens est dans l'existence de groupements sociaux à l'intérieur du milieu où une langue est parlée, c'est-à-dire dans un fait de structure sociale. Il serait assurément chimérique de prétendre expliquer dès maintenant toutes les transformations de sens par ce principe : un grand nombre de faits résisteraient et ne se laisseraient interpréter qu'à l'aide de suppositions arbitraires et souvent forcées ; l'histoire des mots n'est pas assez faite pour qu'on puisse, sur aucun domaine, tenter d'épuiser tous les cas et démontrer qu'ils se ramènent sans aucun reste au principe invoqué, ce qui serait le seul procédé de preuve théoriquement possible ; le plus souvent même ce n'est que par hypothèse qu'on peut tracer la courbe qu'a suivie le sens d'un mot en se transformant. Mais, s'il est vrai qu'un changement de sens ne puisse pas avoir lieu sans être provoqué par une action définie — et c'est le postulat nécessaire de toute théorie solide en sémantique —, le principe invoqué ici est le seul principe connu et imaginable dont l'intervention soit assez puissante pour rendre compte de la plupart des faits observés ; et d'autre part l'hypothèse se vérifie là où les circonstances permettent de suivre les faits de près.

Quelques exemples montreront comment se font les changements de sens et quelle est l'application du principe.

Soit le mot latin *nidus* « nid » ; l'étymologie indo-européenne en est transparente ; c'est le résultat de la combinaison d'un préverbe **ni-* qui marque mouvement de haut en bas et d'un

nom radical **zdo-* appartenant à la racine du verbe latin *sedere* « être établi, être assis » ; ces deux éléments sont de date indo-européenne ; le mot a eu anciennement un sens très vaste, encore conservé dans les langues indo-européennes orientales : en sanskrit et en arménien ; ainsi l'arménien *nist* signifie « lieu où on est établi, résidence, séant » ; mais dans les langues plus occidentales, depuis le slave jusqu'au celtique et au latin, le mot a été limité à un emploi tout particulier — connu même du sanskrit, mais ignoré de l'arménien — celui de lieu où est établi un oiseau « nid » ; les conditions linguistiques du fait sont bien connues : le préverbe **ni-* a cessé d'être employé comme préverbe et ne subsiste plus que dans quelques mots isolés où sa valeur propre n'est plus sensible ; d'autre part, la racine **sed-* n'était plus reconnaissable non plus dans **nizdo-*, et moins encore dans les formes que **nizdo-* a prises dans les diverses langues considérées ; mais on ignore quelles conditions de fait ont pu déterminer la limitation que l'isolement linguistique du mot rendait facile ; comme la limitation est de date indo-européenne, on ne peut faire là-dessus que des hypothèses inconsistantes, et il n'y a lieu que de la constater ici, en attendant que la découverte de quelque fait ou de quelque observation générale permette de déterminer en quel groupe spécial le mot **nizdo-* a pris son sens particulier ; il est permis de conjecturer cependant que ce doit être un terme de chasseur. Du mot latin *nidus* ainsi obtenu, le roman a tiré un dérivé **nidiace(m)* [à l'accusatif], d'où italien *nidiace*, français *niais* ; ce dérivé désigne naturellement « (l'oiseau) au nid » ; il a été employé dans la langue de la fauconnerie, pour désigner l'oiseau pris au nid ; mais pour le fauconnier, l'oiseau au nid est celui qui n'est pas encore dressé et qui est sans habileté : c'est cette dernière notion qui est pour lui la notion dominante. Passant de là dans la langue commune, le mot *nidiace*, *niais* désignera donc un être gauche, maladroit, emprunté, incapable de se tirer d'affaire et ne comprenant rien. Avec la disparition de la fauconnerie, *niais* a perdu toute trace de sens technique, d'autant plus que le rapport linguistique entre *nid* et *niais* n'est plus senti par le sujet parlant français ; *niaiserie*

n'a plus rien de commun avec le sens étroit du mot *niais*, lequel était déjà singulièrement éloigné de celui de la racine **sed-* « être assis ».

Au surplus, la fauconnerie et, d'une manière générale, la chasse, ont fourni beaucoup de termes à la langue commune, on le sait (voir A. Darmesteter, *Vie des mots*, p. 97 et suiv.). Ces divertissements de cercles aristocratiques ayant un prestige particulier, on s'est plu à employer les mots des langues spéciales de ces sports, comme de tous les sports, et on n'a pas tardé à en oublier la valeur exacte ; *leurre* (et *déluré*) sont aussi des termes de fauconnerie qui ont reçu une signification étendue. Le mot *chasser* lui-même est un exemple remarquable, puisque le mot latin vulgaire *captiare*, dont il est la continuation, se rattache à *capere* « prendre » et n'a pu recevoir sa signification particulière que dans la langue des chasseurs. Passé de la langue des chasseurs dans la langue commune, il signifie « pousser devant soi pour prendre » ; en perdant sa précision technique, il passe au sens de « pousser devant soi », et par suite « mettre dehors », si bien qu'un mot dont le sens était « tenter de prendre » aboutit au sens d' « éloigner ». Ici encore, les deux moments, celui de la langue particulière, et celui de la langue commune, se laissent bien distinguer.

On a souvent dit que les langues étaient pleines de métaphores usées. Wundt a déjà montré ce que cette manière de voir a de peu précis et même d'inexact au point de vue proprement psychologique. On voit maintenant, d'un autre point de vue, combien peu on a ainsi une idée des procès réels auxquels sont dus les changements de sens. *Arriver* signifie étymologiquement « aborder », c'est *ad-ripare*, et ce sens s'est bien maintenu par exemple dans le portugais *arribar* ; mais pour un marin, aborder c'est être au terme du voyage : si, de la langue des marins, le terme passe à la langue commune, il signifie simplement ce que signifie le français *arriver*. Le mot *arracher* représente un ancien *ex-radicare* « tirer la racine » ; dans le langage des cultivateurs, ce terme est d'usage fréquent et employé en quantité de circonstances ; s'il passe à la langue commune, la

notion de racine disparaît, et il ne reste que l'idée de tirer un objet engagé dans quelque chose. Le mot *équiper*, emprunté à la langue des marins de la côte normande ou picarde, signifie « pourvoir un bateau de ce qui est nécessaire », et, comme dans la langue technique, l'idée de bateau va de soi, « pourvoir de ce qui est nécessaire » ; que le mot passe dans. la langue commune, et *équiper* n'aura que ce dernier sens : on dit dès lors *équipage* (être en piteux équipage), *équipement* (équipement militaire), sans qu'il reste trace du fait que le centre étymologique du terme est un mot germanique signifiant « bateau », le nom qui subsiste dans anglais *ship*, allemand *schiff*. Dans ces cas et dans les cas innombrables de ce genre, il n'est pas légitime de parler de figures, de métaphores, car tant que les mots sont restés dans la langue particulière, il n'y a pas eu figure à proprement parler, mais emploi d'une manière de s'exprimer où l'idée étymologique n'arrivait pas à la pleine conscience : pour un marin qui aborde, l'idée de rive va de soi, l'essentiel est qu'il arrive au but ; et quand les mots passent de la langue spéciale à la langue commune, ils y passent non avec une valeur étymologique qu'ils ont perdue, mais avec la valeur secondaire qu'ils ont acquise : l'idée d'arriver au rivage qui, pour un marin, subsiste obscurément dans *arriver* est alors éliminée sans même qu'on y prenne garde, car elle n'était plus aperçue.

Ce n'est pas à dire que l'emploi de ces termes empruntés à des langues particulières n'ait pas pour objet de donner à l'expression plus de force et de vivacité : la satisfaction qu'éprouve un marin à parvenir au rivage donnait au mot *arriver* une force de sens qui manquait naturellement au mot de la langue commune ; même sans qu'on se représente d'une manière quelconque le détail d'un échouement de bateau, *échouer*, pris aussi à la langue des marins, exprime l'idée qu'on n'aboutit pas avec plus d'énergie que « ne pas réussir ». Le long usage affaiblit la valeur des mots, et l'emprunt aux langues particulières permet de substituer à des termes inexpressifs des termes auxquels sont associés des sentiments plus vifs. Mais ceci ne fournit qu'un motif pour faire emprunter les mots des langues spéciales, et l'on n'a pas à y chercher un procès indépendant de changement de sens.

Les faits de ce type sont si naturels qu'on les voit se reproduire d'une façon indépendante dans des langues diverses et à des époques diverses. En germanique, une expression composée signifiant « qui (mange) le pain avec un autre » a pris, évidemment dans des groupements militaires, le sens de « compagnon », gotique *ga-hlaiba*, vieux haut-allemand *galeipo* ; sous l'influence germanique, cette expression a été exactement traduite en roman ; de là en français *compain* (cas sujet, aujourd'hui inusité, sauf la forme familière abrégée *copain*) et *compagnon* (cas régime), italien *compagno*, etc. ; nulle part l'idée de « compagnon » n'a un sens plus saisissable que dans un corps de troupes, et l'on conçoit que la langue commune ait trouvé ainsi dans la langue militaire, une manière d'exprimer très fortement cette idée ; mais du fait même du passage dans la langue commune, la notion de partage du pain, qui avait assurément cessé de dominer déjà dans la langue militaire, disparaissait entièrement, et en effet *compagnon* indique simplement en français celui qui est en société intime avec quelqu'un ; de là le mot a de nouveau passé dans une langue particulière, celle des artisans, où il désigne l'ouvrier qui n'est pas patron et qui travaille pour un patron. Un composé tout à fait analogue au gotique *gahlaiba* a été formé, peut-être indépendamment, en arménien ancien où *ənkėr*, littéralement « qui mange avec », signifie simplement « compagnon » ; les rapports des Gots avec les Arméniens n'ont pas été assez intimes pour qu'on soit autorisé à voir dans l'expression arménienne un calque de l'expression germanique ; cependant, la chose est possible à la rigueur. — Au fond, ces expressions remontent à la pratique des repas communiels, pratique religieuse, où se manifestait l'unité du groupe social.

D'autre part, on voit, beaucoup plus tard, la langue militaire fournir un nouveau terme ayant ce même sens, *camarade ; camarade* est l'espagnol *camarada* « chambrée » ; ma *chambrée* a désigné « un compagnon de chambrée » ; et, la notion de compagnon dominant, le mot a passé au sens de « compagnon » dans la langue française commune, et non seulement en français, mais aussi dans les langues voisines, notamment en allemand.

Le mot *camarade* ne signifie pas autre chose que ce que signifie *compagnon* ; mais, surtout au moment de l'emprunt à la langue militaire, il avait une force et une fraîcheur de sens que l'emploi dans la langue commune a fait promptement disparaître.

Parfois le point de départ de tout un développement est un simple nom propre qui prend une valeur déterminée dans un groupe particulier. Ainsi en 1880, un propriétaire d'Irlande nommé *Boycott*, a été, au cours du mouvement nationaliste irlandais, mis à l'index par les Irlandais voisins qui ont refusé d'avoir aucun rapport avec lui ; le procédé a été d'après lui appelé *to boycott* dans le parti nationaliste irlandais ; ce mot a été emprunté par la langue anglaise commune à laquelle l'ont emprunté ensuite les autres langues d'Europe ; quand on emploie l'anglais *to boycott* et, à plus forte raison l'allemand *boycotten*, le français *boycotter*, personne ne pense plus au personnage qui a fourni son nom à ce procédé de lutte sociale et politique, et il va sans dire que le mot désigne quelque chose de beaucoup plus vague et plus général que ce qu'il désignait dans le milieu rural où il a été créé.

Le fait que le mot est pris à une langue particulière n'est pas moins clair en ce qui concerne l'allemand *philister* au sens de « personne fermée aux choses de l'esprit » (et le français *philistin* qui a reçu le même sens sur le modèle de l'allemand, au cours du XIX^e^ siècle). C'est seulement dans des corps d'étudiants, et particulièrement d'étudiants en théologie, que le nom de l'ennemi national du peuple élu, celui des Philistins, a pu, par une association naturelle, devenir celui de tout le vulgaire, par opposition à un petit groupe d'élus, ce qui est arrivé au XVIII^e^ siècle ; aussitôt que le mot s'est étendu aux étudiants des autres facultés, il a pris une valeur plus vague et plus générale, et le sens propre de « peuple des Philistins » s'est effacé de plus en plus ; enfin, en passant dans la langue commune, *philistin* s'est dépouillé presque de toute association avec le nom biblique et, quand Schumann a écrit pour son *Carnaval* une marche des alliés de David contre les Philistins, il réveillait par plaisanterie un vieux souvenir effacé bien plus qu'il ne faisait allusion à une chose courante.

On reconnaît souvent la trace des groupes particuliers où les mots ont séjourné aux nuances de sens qu'ils ont prises. Un mot comme *maréchal*, par exemple, a des sens divers suivant le groupe social auquel la langue générale l'a pris. C'est un mot germanique, venu dans le monde romain par le groupe des conquérants militaires ; le germanique *marahskalk* est un mot composé qui signifiait « garçon de cheval, garçon d'écurie » ; suivant que le mot était employé dans le langage de la cour à désigner un personnage de la suite royale chargé de la surveillance des chevaux ou dans la langue des garçons d'écurie proprement dits, il a pris deux sens bien distincts ; le *maréchal* est un haut fonctionnaire du roi, ou bien il est chargé du soin matériel des chevaux ; et c'est ainsi qu'on a d'un côté le *maréchal* de France, de l'autre le *maréchal* ferrant, ou le *maréchal* des logis, simple sous-officier. — Le *comes stabuli* n'existait qu'à la cour ; c'est le *connétable* ; et en effet le mot latin *comes* n'a persisté que dans le monde de la cour royale perdant ainsi le sens de « compagnon » pour prendre le sens particulier de compagnon du roi, dans le vieux français *cuens* (cas sujet), *comte* (cas régime) ; l'idée de « compagnon » était exprimée par un mot nouveau comme on l'a vu ci-dessus, p. 261.

D'une manière plus générale, le caractère des groupes spéciaux qui ont parlé à un certain moment la langue commune détermine le caractère des innovations sémantiques ; les développements de sens qui se produisent dans les couches inférieures d'une population divisée en classes distinctes ne sont pas les mêmes que ceux qui se produisent dans les couches supérieures. Ainsi l'expression de « voler » ou de « pleurer bruyamment » par *marauder* (agir en matou) qui a été signalée ci-dessus, p. 238, est essentiellement populaire. Le français est la continuation, non du latin classique et littéraire qui a disparu dans la ruine de la civilisation romaine, ni non plus de la langue des chefs germaniques qui ont dominé la Gaule à l'époque mérovingienne et à l'époque carolingienne, et qui était un parler germanique, mais du latin des couches inférieures de la population ; il résulte de là que des éléments de vocabulaire appartenant à la langue du bas

peuple ont remplacé des mots latins communs : « *Caballus* (la rosse) remplace *equus*, *minare* (mener des troupeaux à force de cris) se substitue à *ducere*. Dans la seule désignation du corps humain abondent des exemples de ce genre : *bucca* (la joue gonflée) remplace *os*, *pellis* (peau d'animal) remplace *cutis* ; *perna* (le jambon) ou *camba* (l'articulation entre le sabot et la patte du cheval) remplacent *crus* » (Brunot, *Histoire de la langue française*, I, 131). C'est pour cela que toute la partie un peu relevée du vocabulaire des langues romanes est empruntée, et presque tout entière à la langue écrite.

L'étroitesse du point de départ de certains mots qui sont devenus courants dans la langue commune est parfois surprenante. Ainsi le vieux nom indo-européen du « foie », fidèlement conservé par le latin *iecur*, a disparu de toutes les langues romanes au profit d'un mot de la langue des cuisiniers, mot formé lui-même sur un modèle grec qui a subi en roman même des influences diverses de ce mot grec, à savoir *ficatum* « (foie) garni de figues » ; le nom d'un mets tout particulier est devenu le nom d'un organe (voir Grammont, *Revue des langues romanes*, année 1901, p. 186, avec l'article de G. Paris, auquel il renvoie). De même on a souvent supposé que la « truie farcie » *sus troianus*, ou simplement *troia* (par allusion au cheval de Troie) est devenu le nom de la femelle du porc ; l'hypothèse de l'emprunt à une langue technique est le seul moyen de sauver cette étymologie contestée. Ce sont là des exemples extrêmes, mais qui, par leur caractère excessif même, mettent en pleine évidence quels changements de sens subissent des mots en passant d'un milieu social dans un autre.

Les dictionnaires étymologiques qu'on possède actuellement laissent presque tout à désirer dans l'indication de ces causes de changement. On sait depuis longtemps que le verbe italique qui signifie « dire », latin *dicere*, osque *deicum*, dont une trace se retrouve sans doute en irlandais, est apparenté à une grande famille de mots[1] dont le sens général est « montrer, indiquer »,

1. Les mots rapprochés ici et dans la suite de ce paragraphe pourront sem-

celle de grec δείκνυμι de sanskrit *diçati*, de vieux haut-allemand *zeigôn* (all. mod. *zeigen*). Mais on ne marque pas par quel rapport historique *dicere* se rattache au sens de « montrer, indiquer ». Ce n'est naturellement pas par un rétrécissement abstrait du sens général de « montrer » au sens spécial de « dire », qui n'en est en effet, au point de vue logique, qu'un cas particulier. Le changement s'est produit sans doute de la manière suivante. La racine **deik-* « montrer, indiquer » s'employait en indo-européen avec une valeur juridique définie : à côté de δείκνυμι qui a un sens général, le grec a δίκη qui désigne « l'accusation, le jugement », à côté de *zeigôn*, le vieux haut-allemand a *zîhan* « accuser », *in-zicht* « accusation » ; et le latin même a *index* « celui qui dit le droit », *uin-dex*, *caussi-dicus*, etc., les mots *dicio*, *condicio* sont des termes juridiques, *dicare* indique une proclamation faite dans des formes juridiques ou religieuses définies, et le sens est encore plus net dans *de-dicare* ; l'ombrien *tikamne* (c'est-à-dire *dikamne*) signifie « par consécration ». C'est dans la langue de la procédure, dans *ius dicere* « indiquer le droit » par exemple, que *dicere* a paru avoir le sens de « dire » ; mais ce n'est qu'en passant de la langue juridique à la langue commune que *dicere* a fixé le sens général de « dire » ; du reste ce verbe est resté affecté à tout ce qui se dit dans des formes fixes, et, notamment à la parole publique, et, ainsi que le marquent Bréal et Bailly dans leur *Dictionnaire étymologique latin*, où les emplois juridiques du groupe de *dicere* sont d'ailleurs soigneusement notés, *dicere* est resté le terme solennel qui s'oppose à la causerie désignée par *loqui*. Cet emprunt fait par la langue commune à la langue juridique et religieuse n'est pas un fait isolé ; car la racine **kens-*, celle qui a fourni au latin *censere*, et qui, d'après le témoignage concordant de l'indo-iranien et du latin, avait le sens de « prononcer une formule religieuse ou juridique » a donné au slave et à l'albanais des mots qui signi-

bler assez distants les uns des autres aux personnes qui ne sont pas familières avec la grammaire comparée des langues indo-européennes. En réalité tous ces rapprochements sont rigoureusement justifiés par les règles générales de la phonétique et de la formation des mots, tant en indo-européen commun que dans les langues considérées.

fient purement et simplement « dire ». Il est curieux que la famille du mot slave qui tient exactement la place de la racine **deik-* pour le sens, celle de *kazati* « montrer », ait fourni aussi au russe un verbe dont le sens le plus anciennement attesté est « prêcher », mais qui aujourd'hui signifie « dire », à savoir *s-kazat'*. On n'a pas toujours le moyen de déterminer avec quelque probabilité la série successive des emprunts par lesquels le sens des mots s'est progressivement transformé, mais des exemples tels que celui-ci indiquent au moins en quelle direction on a chance d'apercevoir l'explication des changements qu'on constate d'une période linguistique à une autre.

Toutefois on ne saurait démêler les actions et réactions complexes auxquelles sont dus les changements de sens, là où l'histoire des faits n'est pas exactement connue. Sans des témoignages historiques détaillés on n'aurait sans doute pu arriver à reconnaître comment, c'est-à-dire dans quels groupes sociaux, un mot qui signifiait en latin « tambour » a pu prendre en français le sens de « timbre-poste » (voir A. Darmesteter, *Vie des mots*, p. 81 et suiv., sur le mot *timbre*). Par le fait même qu'ils dépendent immédiatement de causes extérieures à la langue, les changements sémantiques ne se laissent pas restituer par des hypothèses proprement linguistiques.

Il est dès lors impossible, on l'a vu, de donner une démonstration en règle de la théorie proposée ici ; cette démonstration ne pourrait résulter que de l'examen de tous les changements de sens constatés dans une langue donnée entre deux périodes données et de la constatation que tout ce qui ne s'explique pas par des causes proprement linguistiques ou par des changements des choses désignées provient du passage des mots de langues particulières à la langue commune ou du passage inverse de la langue commune à une langue particulière ; pareille constatation est irréalisable en l'état actuel des connaissances ; car on n'a, sur aucun domaine linguistique, le moyen de procéder à un examen complet de cette sorte. Mais là même où aucune indication de fait ne permet de marquer par quelle série d'emprunts intérieurs un mot a changé de sens, la possibilité de ces passages

demeure vraisemblable la plupart du temps, et on est obligé de les supposer si l'on ne veut pas admettre que les générations successives ont, par pur caprice, associé des notions différentes à un seul et même mot. Les conditions psychiques de la sémantique sont constantes ; elles sont les mêmes dans les diverses langues et aux diverses périodes d'une même langue ; si donc on veut expliquer la variation, il faut introduire la considération d'un élément variable lui-même, et, étant données les conditions du langage, cet élément ne peut être que la structure de la société où est parlée la langue considérée.

IV

Ces principes une fois posés, la méthode qu'il convient d'appliquer dans l'étude de la sémantique ressort, semble-t-il, assez nettement. En présence d'un mot donné il convient d'examiner tout d'abord la forme du mot et son degré d'isolement dans la langue ; un mot isolé se comporte autrement qu'un mot qui fait partie d'un groupe ; il faut d'autre part se rendre compte de l'influence possible de la forme, du rôle dans la phrase des associations phoniques qu'il éveille (voir Grammont, *Onomatopees et mots expressifs*, *Revue des langues romanes*, XLIV, 97 et suiv.). En second lieu, on doit suivre l'histoire des choses signifiées, qui réagit sur le mot et sur ses connexions avec le reste du vocabulaire. Enfin, et surtout, il faut marquer par quels groupes sociaux le mot a été transmis, passant d'une langue particulière à une autre langue particulière. Ce sont là autant de procès distincts, que l'analyse doit isoler, car ils sont d'espèces différentes : mais dans la réalité ces diverses actions ne se séparent pas les unes des autres ; elles s'appliquent à un même mot tantôt simultanément et tantôt successivement ; elles se combinent de telle sorte qu'il devient souvent malaisé de marquer ce qui revient à chacune d'elles. De plus les passages de la langue commune aux langues particulières sont en grande partie insaisissables ; car il n'y a nulle part de limite précise entre les langues particulières et la langue commune, et dans la

mesure où la différence se laisse préciser, il y a réaction constante du vocabulaire commun sur les vocabulaires particuliers et des vocabulaires particuliers sur le vocabulaire commun : ce n'est que par abstraction qu'on a pu ci-dessus isoler le passage du mot de la langue commune dans un vocabulaire particulier ou inversement ; dans la plupart des cas, le passage est incessant, et il y a va et vient du mot entre les deux vocabulaires. Les complications qui résultent du croisement de tous ces faits d'espèces variées sont inextricables d'autant plus que chaque procès comporte un nombre illimité d'actions autonomes de chacune des trois espèces définies au début de cet article. Et de plus ce n'est que par hypothèse qu'on peut apprécier le degré d'influence de chacune des actions qu'on envisage. Enfin il est toujours impossible de faire un dénombrement complet des actions qui interviennent pour un mot donné, car les renseignements que l'on possède ne suffisent jamais pour cela. L'examen d'une question de sémantique se réduit dans la pratique, la plupart du temps, à un examen de possibilités et de probabilités, et il n'est pas licite d'ordinaire d'aboutir à des conclusions trop résolument affirmées ni trop absolues.

Le mot haut-allemand *rappe* signifiait à l'origine « corbeau » : c'est la forme du haut-allemand qui répond exactement à l'allemand classique *rabe* ; cette forme a été employée au moyen âge pour désigner une monnaie où figurait la tête de corbeau qui se trouve dans les armes de la ville de Fribourg-en-Brisgau ; plus tard on a employé la même forme pour désigner un « cheval noir » ; si la forme était demeurée confinée en haut-allemand, elle aurait eu peine à se fixer en ce sens, mais elle a été empruntée vers le XVII[e] siècle par les autres dialectes allemands, sans doute dans les armées de mercenaires qui étaient composées d'hommes de provenances diverses ; *rappe* n'avait pas dans ces groupes son sens de « corbeau » et signifiait seulement « cheval noir ». C'est en ce sens que l'allemand moderne a adopté ce terme, et il en est résulté que dès lors les dialectes même du haut-allemand ont tendu à ôter ce sens de « corbeau » à *rappe* et à n'employer pour désigner l'oiseau que le mot commun *rabe*.

Le latin *augur* désigne le citoyen qui était officiellement chargé d'examiner le vol des oiseaux et de tirer de là des conséquences sur l'issue des affaires publiques en discussion. Au point de vue étymologique, c'est un nom composé dont le premier terme est le thème de *auis* « oiseau », mutilé par une altération phonétique ; le second terme n'était pas reconnaissable en latin même, si bien que, aujourd'hui encore, on n'en saurait déterminer la nature et le sens d'une manière certaine ; au point de vue latin le mot n'est donc pas clair pour la forme, et a l'aspect d'un mot isolé. Les dérivés *augurium* « prévision de l'avenir par l'auguration » et *auguror* « je prévois l'avenir par l'auguration » ont pris dans la langue spéciale des magistrats romains, le sens de « prévision de l'avenir » qui était la représentation dominante à leur point de vue, l'auguration n'étant qu'un moyen. Dès lors la langue commune a été amenée à employer ce mot pour toute prévision de l'avenir, au moins quand on voulait s'exprimer d'une manière noble ; déjà dans la tragédie latine *auguror* est employé pour signifier « je m'attends à » ; plus l'auguration devenait un simple rite auquel on attribuait moins de foi, et plus ces mots se limitaient au sens de prévision de l'avenir que le triomphe du christianisme rendait enfin le seul possible : c'est celui que présentent les mots romans issus de *auguror* (prononcé *aguror* en latin vulgaire), par exemple espagnol *agorar* signifie « prévoir, s'attendre à », les mots *bonum agurium*, *malum agurium* ont abouti de leur côté à français *boneür*, *maleür*, d'où *bonheur*, *malheur*, qui, entrant tout à fait dans la langue commune, n'ont pas gardé trace même du sens ancien de « prévision », « attente de l'avenir » : le mot (*h*)*eur*, issu d'*agurium*, a pris à lui seul le sens de *bonheur* par opposition à *malum agurium*, d'où le dérivé *heureux*, qu'on applique à tout événement agréable, à toute personne favorisée du sort, et même à tout ce qui est réussi. Entièrement séparé de son sens étymologique par des circonstances linguistiques, puis historiques, *augurium* a abouti aussi en français à un sens très vague par suite du passage du mot de la langue des magistrats romains dans celle de cercles de plus en plus étendus.

Le mot *hospitale* « lieu où l'on reçoit des hôtes » s'est trouvé séparé du mot *hospes* dont il est dérivé en latin ; en effet le suffixe *-ale* a cessé d'être productif; il n'y a donc presque plus rien eu de commun en français ancien entre *oste* et *ostel* ; le mot *ostel* a été appliqué dans certains groupes d'individus à désigner la grande maison où ils recevaient l'hospitalité, où ils étaient hébergés ; cette grande maison peut être suivant les cas un *hôtel-dieu*, où l'on reçoit des malades et des infirmes (c'est aussi le sens de l'italien *ospedale*, *spedale*), une maison de ville ou *hôtel* de ville, un *hôtel* de voyageurs, ou une grande maison particulière. De ces sens particuliers, deux ont survécu et ont passé dans la langue commune en s'isolant de plus en plus l'un de l'autre, celui d'*hôtel* de voyageurs, qui a fourni les dérivés *hôtelier*, *hôtellerie*, etc., et celui d'*hôtel* particulier ; en ce dernier cas, ce mot a pris dans la bourgeoisie parisienne du XIX^e^ siècle un sens tout particulier : celui de maison séparée consacrée à une seule famille, par opposition aux maisons de rapport, divisées en appartements séparés, et louées à des locataires différents ; et dès lors, on a pu habiter un *petit hôtel*, c'est-à-dire une petite maison séparée. Sauf la forme isolée *hôtel-dieu*, *hôtel* ne sert plus à désigner l' « endroit où l'on reçoit des malades, des infirmes » pour une raison historique : les maisons de cette sorte étaient essentiellement des fondations pieuses, et le nom qui a prévalu est la forme latine savante provenant de la langue du clergé, *hospital*, d'où *hôpital* ; pour la même raison, le mot *hospitium* a été aussi employé à un usage analogue sous la forme légèrement francisée *hospice* ; et il y a eu répartition du sens en français de Paris entre l'*hôpital*, qui reçoit les malades, et l'*hospice*, lieu de refuge pour les infirmes et les vieillards. Ces mots, et surtout le mot *hôpital*, sont à leur tour entrés dans la langue commune avec l'importance prise par cette forme d'assistance dans la vie parisienne ; rien n'y indique plus la notion de réception d'un hôte, et la représentation qui domine est celle de soins à donner à des malades.

Le mot grec ἐκκλησία, de la famille de ἐκκαλέω « j'appelle, je convoque » signifiait « assemblée » en grec ; dans les milieux

chrétiens, il a désigné spécialement l'assemblée des fidèles : il a passé en ce sens dans la langue spéciale des chrétiens de Rome ; là il a désigné l'assemblée des chrétiens (voir Kretschmer dans *Zeitschrift für vergleichende Sprachforschung*, XXXIX, 539 et suiv.) ; d'autre part le mot ἐκκλησία signifiait « lieu de réunion des fidèles », exactement comme *marché* signifie « lieu où on tient le marché » ; le développement de sens se produit tout naturellement dans des phrases comme « je vais au marché », « je vais à la réunion » : le latin a pris aussi le mot grec en ce sens ; comme, en latin, le sens de réunion, convocation, n'étaient pas attachés au mot, isolé de toutes ses connexions linguistiques par l'emprunt, et que *ecclesia* était un pur terme de langue particulière, sans usage dans la langue commune, ces deux sens de « groupe des fidèles » et de « lieu de réunion des fidèles » se sont fixés sans aucun mélange, et ils se sont transmis aux langues romanes ou du moins au groupe occidental des langues romanes. Sur le sol français, le mot *église* est entré dans la langue commune, avec le christianisme qui est devenu la religion de tous les habitants du pays ; les fidèles ont cessé de former une « assemblée » pour devenir un vaste groupe uni par une foi commune et par des institutions communes, et l'*ecclesia*, devenue en français l'*église*, a été le nom de ce vaste groupe en même temps que des bâtiments où se réunissaient les fidèles ; dans la langue populaire, qui envisage surtout des choses, il ne désigne même guère que les bâtiments.

Ces exemples, où l'on a remarqué seulement les plus gros faits et les plus généraux, permettent de se faire une idée de la manière dont les faits linguistiques, les faits historiques et les faits sociaux s'unissent, agissent et réagissent pour transformer le sens des mots ; on voit que, partout, le moment essentiel est le passage d'un mot de la langue générale à une langue particulière, ou le fait inverse, ou tous les deux, et que, par suite, les changements de sens doivent être considérés comme ayant pour condition principale la différenciation des éléments qui constituent les sociétés.

LE NOM DE L'HOMME [1]

L'allemand a deux noms de l'« homme » : l'un désigne le « mâle », l'individu capable de porter les armes, l'autre est le nom générique de l'espèce humaine : *mann* et *mensch* ; de même le néerlandais. Le second de ces noms est un dérivé du premier. Ni en français ni en anglais, on n'observe rien de pareil : *homme* et *man* servent à la fois pour l'une et pour l'autre notions. L'état de choses allemand concorde avec celui de la plupart des anciennes langues indo-européennes : le sanskrit oppose *mánuṣaḥ* à *vīráh* et à *nā* (accusatif *náram*), le grec *ánthrōpos* (dont l'étymologie est obscure) à *anêr*, l'arménien *mard* (dont le sens premier est « mortel ») à *ayr* (qui répond à grec *anêr*), le lituanien *žmū̃* et *žmogùs* à *výras*, le slave *člověkŭ* à *mǫžĭ*, le gotique *guma* à *wair*, l'irlandais *duine* à *fer*, le latin enfin *homō* à *uir*.

Il est curieux que cette distinction qui paraît si naturelle, et qui est si commode, ait tendu à s'effacer. Le latin, le germanique la possédaient; les langues romanes, l'anglais l'ont perdue. Il semble que la distinction ait tendu à s'éliminer durant la période où les langues passaient par la période plébéienne : on sait que les langues romanes continuent, non le latin des hautes classes de la société, mais un latin plébéien, comme on le voit et par bien des détails de vocabulaire, et par certains caractères de leur grammaire. Le nom noble de l'« homme » a disparu ; il n'est resté que le nom générique : *homo*. De même en anglais : l'anglais a passé par une période toute plébéienne lors de la conquête

1. D'après une conférence faite au Cercle français de l'Université d'Amsterdam.

franco-normande, à la suite de laquelle les classes dominantes ont parlé français.

Si le latin *uir* a disparu des langues romanes, il en est resté un dérivé, mais dont, dès l'époque romaine, le lien avec *uir* était si relâché qu'on cessait d'unir les deux mots: *uirtūs*. La *uirtūs*, c'est l'ensemble des qualités qui font un « mâle », un « guerrier »: *appellata est ex uiro uirtus*, dit Cicéron dans les Tusculanes. C'est ce que le grec attique nommait *andreíā*, la langue homérique *ēnorée̅*, la qualité de l'*anēr*. Mais, en latin, le mot a été affecté à désigner toute qualité morale, ce que le grec nommait *aretḗ*: faute d'avoir un mot propre, on s'est servi de *uirtūs* pour rendre le mot grec, et ainsi le mot latin s'est empli d'un sens nouveau, qui le séparait définitivement de *uir*; avec le christianisme, ce sens a pris une nuance nouvelle qui l'a isolé plus encore de *uir*. D'autre part, *uirtūs*, qui désignait un mérite actif, a servi à désigner les propriétés actives d'un produit. Le français qui n'a aucun représentant de *uir*, a conservé ainsi *uirtūs*, sous la forme *vertu*, avec la double valeur de « qualité morale » et de « propriété active (d'un objet) ». On voit ici comment les classes sociales qui créent la culture agissent sur le vocabulaire. Du reste, ce n'est plus guère que dans la langue littéraire et savante que *vertu* a gardé sa valeur. La prédominance des préoccupations matérielles, le souci d'une lutte âpre pour les meilleures situations sociales ont enlevé à la « vertu » beaucoup de son prestige. Dans la langue courante, un peu populaire, le mot « vertu » n'existe plus que dans des emplois ironiques : *il faut de la vertu pour cela, il en a de la vertu !* Ainsi le mot qui désignait le mérite de l'homme fort ne sert plus dans le français familier d'aujourd'hui qu'à indiquer la naïveté de l'homme qui est trop « bon » pour son siècle, et qui par suite est la dupe des autres.

*
* *

Le nom générique de l' « homme » en latin est *homō*. Le mot est ancien ; on en est averti dès l'abord par une alternance voca-

lique : à côté de *homō*, il y a une forme à vocalisme radical *e*, *hemonem*, attestée par Festus, et que d'ailleurs *nēmō*, c'est-à-dire **ne-hemō* « pas un homme » suffirait à indiquer. On enseigne souvent que l'*o* de *homō* serait une altération phonétique de l'*e* de l'áncien *hemō* ; c'est une erreur : d'abord cette altération n'entrerait dans aucune règle connue ; et surtout les autres langues du groupe italique ont aussi *o*, osque *humuns* valant *homines*, ombrien *homonus* valant *hominibus*, sans que rien y indique un passage de *e* à *o*. Des noms correspondants de l' « homme » se retrouvent, et encore avec un autre vocalisme, qui n'est ni *e* ni *o*, mais zéro, dans gotique *guma*, vieux haut allemand *gomo* (le mot, disparu aujourd'hui à l'état isolé, est celui qu'on a à la fin de *bräuti-gam*), et dans lituanien *žmů*, vieux prussien *smoy*, toujours avec le même sens.

Ni en latin, ni dans l'ensemble des parlers italiques, ni en germanique, ni dans les langues baltiques, le mot ne se laisse interpréter par d'autres mots existant dans la langue. Le nom de l' « homme » n'y a pas plus de sens que le nom *equos* du « cheval », le nom *ouis* de la « brebis » n'en ont en latin ; ce sont des désignations qui ne se laissent pas expliquer. Toutefois, on aperçoit tout de suite une différence : bien que la formation soit la même dans toutes les langues et que, par exemple, la flexion gotique de *guma*, *gumins* soit parallèle à la flexion latine de *homō*, *hominis*, le vocalisme radical diffère d'une langue à l'autre : latin *homō* et *hemō*, en face de gotique *guma* et lituanien *žmů*. Ceci laisse entrevoir que le nom de l' « homme » serait le dérivé d'un mot plus simple et que, tout en étant parallèles, les formations italique, germanique et baltique auraient été fixées indépendamment les unes des autres.

En effet, au point de vue indo-européen, ces mots se laissent interpréter aisément. La comparaison des données indo-iraniennes et grecques les plus anciennes et les traces que les conceptions ont laissées dans le vocabulaire de plusieurs langues permettent de se rendre compte, en quelque mesure, de la manière dont, dans le monde indo-européen, se comprenait l'opposition de « homme » et de « dieu ». Les formules homériques sont

claires : les hommes sont « mortels », *brotoí*, *thnētoí*, et « terrestres », *epi-khthónioi* (θ 479 par exemple) ; les dieux sont « immortels », *ámbrotoi*, *athánatoi*, et « célestes », *ep-ouránioi*. Quand, au chant VI de l'Odyssée, Ulysse adresse à Nausicaa son joli discours, il lui dit : « Es-tu une déesse ou une mortelle ? Si tu es de ces déesses qui habitent le ciel, tu ressembles de près à Artémis, fille du grand Zeus, par l'aspect, par la grandeur, par la taille ; si tu es de ces mortels qui habitent sur la terre, trois fois heureux ton père et ta noble mère, trois fois heureux tes frères. » (ζ 149-155.)

De préférence au vieux mot qui subsiste en sanskrit dans *mánuḥ* et *mánuṣaḥ*, en germanique dans gotique *manna*, etc., en slave, dans le dérivé *mǫžĭ*, on a désigné souvent l' « homme » en général par ces épithètes qui l'opposent aux dieux. Chez Homère, les hommes sont sans cesse appelés « mortels » : *brotoí* ou *thnētoí* ; le nom iranien de l'homme est couramment un dérivé qui signifiait « mortel », *martiya-* en vieux perse, *mard* en persan ; de même, en arménien, le nom de l'homme est *mard*, qu'on n'a aucune raison de tenir pour emprunté à l'iranien. On doit donc s'attendre à trouver ailleurs l' « homme » désigné par son épithète de « terrestre », qui l'oppose aux dieux « célestes ».

Or, le principal des noms indo-européens de la « terre », celui qui est le plus sûrement ancien et le plus général, est précisément le mot dont lat. *homō*, etc. sont des dérivés. Ces noms de l' « homme » ont à l'origine signifié « terrestre », et ce n'est que par oblitération à la fois de la forme ancienne du mot et des conceptions anciennes que ces mots ont pris la valeur toute abstraite qu'on leur voit dans les parlers italiques, germaniques et baltiques.

Ce nom de la « terre » ne s'est pas conservé en germanique. Mais il subsiste en latin dans le dérivé *humus* qui a gardé le genre féminin (on en a des dérivés tels que *humilis*) et en baltique où il est le nom courant de la « terre » : lituanien *žẽmė*, vieux prussien *semme* ; le mot slave correspondant est *zemlja* ; le grec a *khamaí* « sur terre », et avec des formes phonétiques un peu différentes, *khthōn* « terre ». Les correspondants indo-iraniens

sont d'usage courant. Latin *homō* et *hemō*, gotique *guma*, lituanien *žmũ* sont des dérivés du thème **ghem-*, **ghom-*, **ghm-*, qui était en indo-européen le principal nom de la « terre ». Pour la forme, ces dérivés ne se comprennent que si on se reporte à l'indo-européen, où ils sont de type normal. Pour le sens, ils renvoient à un temps où, toute pensée étant de type religieux, il était naturel de désigner l' « homme » par les traits qui le distinguent des dieux : la mortalité, l'habitat sur la terre.

* *

Mais — c'est le fait qui domine tout le développement du vocabulaire, comme toute l'histoire des langues — un mot n'est pas lié par ses origines étymologiques. Une fois créé, il se libère par l'usage qui lui donne son autonomie, par les accidents qui rompent le lien avec le primitif qui a servi à le former, et enfin par le changement des conceptions. Dès qu'un mot comme latin *homō*, *hemō* entre en circulation et qu'il sert couramment à désigner l' « homme », on cesse de penser au primitif dont il est issu, même si la formation en reste claire : le seul fait que *deviner* est un verbe d'usage à peu près courant fait qu'on ne pense plus à *devin* en le prononçant, bien que la formation soit demeurée transparente jusqu'au français actuel. D'ailleurs un Romain ne pouvait apercevoir aucun rapport entre les formes de *homō*, *hemō* et celles de *humus*. Enfin on a cessé, de bonne heure, d'opposer les « hommes » et les « dieux » comme le fait encore Homère si souvent. Il n'y a plus eu qu'un mot isolé, *homō*, *hemō*, qui est le nom générique de l' « homme ».

Un nom de ce genre comporte plusieurs développements de sens possibles.

D'abord il se prête à suivre tous les changements de conception qui interviennent au fur et à mesure que les idées changent : *homō* a servi à traduire gr. *ánthrōpos*, et il a rendu tous les sens que la philosophie grecque a donnés à *ánthrōpos*. Le christianisme, à son tour, lui a donné des nuances nouvelles.

En second lieu, par le fait qu'il désigne l'« homme » d'une

manière quelconque, il est sujet à prendre une valeur indéfinie dans les phrases négatives, interrogatives et conditionnelles. Le fait est intervenu dès le latin ancien pour *nēmō* : l'utilisation de la forme *hemō*, qui n'a pas survécu ailleurs, la disparition de *h* entre voyelles et la contraction ont séparé *nēmō* de *homō*. Mais, à l'époque romane, le développement a recommencé : le nominatif *homō* a pris une valeur indéfinie dans les phrases négatives ou conditionnelles ou interrogatives ; ainsi s'est formé le français *on* où survit le nominatif, tandis que *homme* représente l'accusatif *hominem*. Et, enfin *on*, comme l'allemand *man*, est devenu ce mot si commode dont il semble qu'on ne puisse se passer, et dont cependant se passent tant de langues : *on* ; le français en est arrivé à l'employer comme un pronom universel, qui remplace parfois *tu*, souvent *nous*.

En troisième lieu, si avec le progrès d'un état de choses policé, l'homme par excellence cesse d'être l'homme armé, le soldat, il y a une valeur du type *uir* qui reste à exprimer : c'est celle de l'homme considéré au point de vue du sexe, l' « homme » opposé à la « femme ». C'était une des valeurs de *uir*, et elle est si nécessaire que la langue savante a pris, surtout en ce cas, *uirilis*— français *viril*—au latin. Or, le mot *homō* est du genre masculin ; il n'en peut être autrement parce que le genre masculin n'est pas seulement celui qui désigne les mâles, c'est aussi le genre commun, celui dont on se sert dès qu'on ne spécifie pas qu'il s'agit d'un être de sexe féminin ; le français *homme* a gardé ce genre ; comme le seul sens que possède le masculin, dans les cas exceptionnels où il en a un, est celui d'indiquer le sexe du mâle, le mot *homō* s'est trouvé tout naturellement conduit à indiquer aussi l'homme en tant qu'il est de sexe mâle, le *uir*. Et c'est, des diverses valeurs de *homme*, celle qui a pris le plus d'importance dans la langue populaire. En effet, pour quiconque ne pense pas abstraitement ou religieusement, pour qui ne réfléchit pas à la position de l'homme dans le monde ou par rapport à la divinité, l'emploi du mot « homme » se réduit au type indéterminé : *un homme est venu*. A part cela, ce qui se trouve le plus usuellement, c'est l'opposition de l' « homme » et de la « femme ». Le

résultat est que, pour le sentiment d'un Français du peuple, le mot « homme » désigne avant tout l'opposé de la femme. Pour une femme du peuple, *mon homme* est la désignation constante du mari, comme la *femme* est la désignation universelle de l'épouse, même chez les gens qui parlent une langue distinguée. Rendu possible par une circonstance grammaticale, ce développement de sens est devenu dominant dans l'usage populaire. Le caractère naturel de ce changement ressort de l'état de choses germanique où got. *guma*, v. angl. *guma*, v. h. a. *gomo*, etc. désignent le « *uir* » et non l' « homme » en général ; le *-gam* de l'allemand *bräuti-gam* « fiancé » porte encore aujourd'hui témoignage de cette valeur du mot. Or, la valeur primitive du mot germanique ne pouvait être que celle de latin *homō*.

*
* *

Le développement de sens vers l'indéfini, qui aboutit à *on*, ou celui du sens sexuel, qui a abouti dans la langue populaire à celui de « mari », proviennent l'un et l'autre du fait que, chez les couches peu cultivées de la population dont les parlers connus à date ancienne et le français populaire d'aujourd'hui sont l'expression, il n'y a que rarement lieu d'envisager l'homme à un point de vue général et abstrait. On voit ainsi ce qu'une expression formée dans le langage religieux et plus élevé, comme *homō*, devient dans le parler populaire.

D'une manière générale, le vocabulaire populaire se nourrit en grande partie de termes qui lui sont fournis par la langue savante. On le voit aujourd'hui encore pour l'expression de l'idée d' « homme ».

Dans le parler parisien courant, le mot *type* tend à remplacer *homme*. On dit couramment : *c'est un bon type*, *c'est un sale type*, *il est venu un type que je ne connais pas*. Le mot *type*, par le fait qu'il est de genre masculin, ne s'applique guère qu'à des gens du sexe masculin. Or, c'est un mot on ne peut plus sûrement d'origine grecque. C'est la langue des écoles qui l'a fourni au parler courant.

Un autre mot, presque aussi savant par ses origines, a une valeur plus générale : le mot *monde*. Quelle qu'en soit, en latin, l'origine assez complexe et assez singulière [1], ce terme a pris une valeur nouvelle par le fait qu'il a servi à traduire le grec *kósmos* dont il a pris tous les sens. Il est devenu ainsi le mot qui désigne l'univers organisé. Dans la langue chrétienne, il en est venu à désigner l'univers terrestre, par opposition aux choses célestes ; il a indiqué ce que le mot *siècle* indiquait aussi en partie. Le français *monde* a hérité de toutes ces valeurs diverses, et il se prête à des usages très variés : il sert à désigner l'univers en général, plus particulièrement l'univers en tant qu'il est habité : on dit volontiers *le monde habité*, — les gens qui vivent de la vie du siècle, et surtout ceux de certaines classes, *le grand monde*, *le beau monde*, *le demi-monde*, *le pauvre monde*, — mais encore l'ensemble des hommes, *tout le monde*. On a été amené par là à se servir du mot *monde* pour désigner les « hommes » en général ; et dès lors, on peut arriver à dire populairement (j'ai entendu la phrase en Berry) : *c'est du monde trop grand*, pour dire : « ce sont des gens de trop grande taille ». En ce sens, *le monde* s'oppose nettement aux animaux : on dira qu'il faut garder une certaine nourriture « pour le monde », c'est-à-dire ne pas la donner aux animaux.

L'adjectif qui correspond à *homō* est en latin *hūmānus*. La formation en *-ānus* a son pendant dans *germānus* par exemple. On voit mal comment expliquer l'*ū* de *hūmānus* en face de *homō* ; mais, d'où que sorte l'*ū*, cet adjectif *hūmānus* est à *homō* ce que *dīuīnus* est à *deus* ; le plus probable est qu'il faut partir d'un nominatif **hōm* du nom de la « terre » que le latin n'a pas gardé, et qui passait phonétiquement à **hūm* comme on a *fūr* en face du grec *phôr* ou *cūr* qui représente *quōr*. La formation de *hūmānus* paraît du reste ancienne ; le lituanien a un dérivé *žmo-gùs*, et, au pluriel, *žmónės* « les hommes » ; le vieux prussien, *smonenawins* « l'homme ». Mais, si la formation de *hūmānus* est

1. M. Vendryes a démêlé cette origine dans un curieux article des *Mémoires de la Société de linguistique*, XVIII, 305 et suiv.

ancienne, le mot a pris en latin savant un sens nouveau : Cicéron s'en sert pour traduire *philánthrōpos*, et d'*hūmānitās* pour traduire *philanthōpiā*, c'est-à-dire pour désigner tout ce que comporte la notion d'un homme cultivé, qui a reçu toute la culture de son temps. De par sa nature d'adjectif, le mot *humain* est d'ailleurs plus abstrait et plus général que le nom même de l' « homme ».

Partout et toujours, les variations de sens des mots dépendent des classes sociales qui les emploient : compte tenu de la structure de la langue, changement de sens signifie le plus souvent pour un mot passage d'une classe sociale à une autre.

QUELQUES HYPOTHÈSES
SUR DES
INTERDICTIONS DE VOCABULAIRE
DANS LES LANGUES INDO-EUROPÉENNES[1]

En des points du monde très divers, et dans des langues très variées, on observe des « tabous » de vocabulaire, dont on trouvera un aperçu chez Frazer, *The Golden Bough*, Part II, *Tabou and the Perils of the Soul*, p. 318 et suiv. ; *Rameau d'or*, trad. franç., I, p. 331 et suiv. ; dans l'Afrique du Sud, dans le domaine des langues malayo-polynésiennes (y compris Madagascar, v. van Gennep, *Tabou et totémisme à Madagascar*, p. 104 et suiv.), en Extrême-Orient et aussi en Europe, notamment dans le Nord de l'Europe, il apparaît que certains mots sont interdits par l'usage, soit à un groupe d'hommes, soit à des individus déterminés, soit durant certaines périodes, en certaines occasions ; on « taboue » par exemple le nom d'un mort, celui d'un chef, celui des membres de la famille où l'on prend femme, etc. ; et le tabou ne touche pas seulement les noms propres en question, mais il s'étend aux noms communs, identiques ou non à ces noms, qui sonnent d'une manière identique ou analogue, ou même partiellement analogue.

On a souvent répété que les langues des peuples peu civilisés changeraient avec une grande rapidité et deviendraient méconnaissables au cours d'une seule génération : dans la mesure où elle est exacte, l'observation ne saurait guère se justifier que par des faits de ce genre ; car ces langues ne sont pas sujettes à des changements particulièrement rapides ; elles semblent même

1. Brochure, dédiée à M. J. Vendryes, 3 juillet 1906 (per nozze) ; non mise dans le commerce.

avoir une remarquable stabilité dans beaucoup de cas, en Polynésie par exemple ; mais il peut arriver qu'un voyageur, à un second passage, trouve supprimés par des tabous beaucoup de mots qu'il s'était fait enseigner quelques années auparavant.

A date historique, cet usage est assez peu attesté chez les peuples de langue indo-européenne, ou du moins il affecte des formes nouvelles ; dans la langue polie on évite depuis longtemps les termes qui expriment proprement et précisément un certain nombre de fonctions naturelles ; l'influence de cet usage sur le vocabulaire est bien connue ; on sait par exemple comment le féminin de *gars* a disparu du français ordinaire parce qu'il était devenu le nom propre de la prostituée, et comment *fille* étant venu à être affecté à ce sens, on doit dire *jeune fille* là où l'on disait autrefois *fille*. On sait aussi comment les jurons ont été modifiés pour éviter les blasphèmes interdits et comment le français a *morbleu*, etc., etc., au lieu de *mort Dieu*, etc., et l'anglais *gog*, *cocke*, etc., au lieu de *god*.

Il est permis de supposer que, sous sa forme commune, l'usage du « tabou » a dû exister à date ancienne dans les populations de langue indo-européenne, et ce serait un moyen d'expliquer la perte de certains mots dont on ne s'explique pas aisément la disparition sur une partie du domaine. Le nom de l' « ours » en fournit sans doute un bon exemple.

L'ours était commun sur toute l'aire occupée au début de l'époque historique par les langues indo-européennes (v. O. Keller, *Thiere des classischen Alterthums*, p. 106 et suiv.). Et il en existe une désignation dont le caractère indo-européen est certain :

skr. *ṛkṣaḥ*, zd *aršō* (ἄρκτος obtenu par une correction sûre, v. Bartholomae, *Altiran. Wört.*, s. v.), persan *xirs*, afgh. *yaž*, signī *yūrš'*, sariqolī *yürx*, yidghah *yerš*, māzandarānī *aš*, ossète *ars* (v. W. Miller, *Spr. d. Osseten*, p. 32).

gr. ἄρκτος ; ce mot sert à la fois pour le mâle et pour la femelle; et tel était l'usage indo-européen ; le féminin sanskrit classique *ṛkṣī* est une nouvelle formation nécessitée par le fait que le sanskrit n'a pas conservé de thèmes en *-ā-* féminins ; le lat. *ursa*, dont la formation est différente, est aussi nouveau et remplace

un **ursus* féminin, de même que *lupa* remplace un *lupus* féminin attesté chez Ennius et chez Varron (v. Wölfflin, *Arch. f. lat. Lexicogr.*, III, 562 et VII, 280 ; Delbrück, *Vergl. Synt.*, I, p. 114; Neue-Wagener, *Lat. Formenlehre*, I³, p. 925 ; Niedermann, *Contribut. à la crit. et à l'explicat. des gloses lat.*, p. 35 ; Wackernagel, *Altind. gramm.*, II, p. 17 ; A. Meillet, *Études sur l'étym. et le vocab. du v. sl.*, p. 246).

lat. *ursus* ;

irl. *art*, gall. *arth* ; la dédicace *deae artioni* « à la déesse à l'ourse », inscrite sur une statue trouvée dans la région bernoise, garantit aussi l'existence du mot en gaulois (S. Reinach, *Rev. Celt.*, XXI, p. 288 et suiv. ; Zupitza, K. Z., XXXVII, p. 393 n. ; Dottin, *Manuel p. s. à l'ét. de l'ant. celtique*, p. 240) ; sur le *t* du mot celtique, v. *Mémoires Soc. Linguistique*, XI, p. 316 et suiv. (et cf. Pedersen, K.Z., XXXVIII, 208 n.) ;

alb. *ari* ;

arm. *arj* (thème en *-o-*, gén. *arjoy*) ; le *j* fait difficulté (v. M. S. L., X, 281 n., et Pedersen, K. Z., XXXVIII, 208 et 213; XXXIX, 432), mais il n'y a pas de raison valable d'écarter le mot arménien, comme le fait M. Scheftelowitz, BB., XXVIII, 293 ; dans la traduction arménienne de son travail, M. Pedersen maintient le rapprochement de arm. *arj*, et ne juge pas utile de mentionner les objections de M. Scheftelowitz (revue arménienne *Hantes*, mai 1906).

Il existe un doublet à consonne intérieure simple du même mot :

gr. ἄρκος, ἀρκίλος, et sans doute le nom propre de peuple Ἀρκάδες ;

Le doublet gr. κτ : κ est d'un type indo-européen (cf. gr. χθον- : χαμ-, skr. *kṣam-* : zend *zəm-* « terre ») et confirme par suite l'antiquité du nom de l'« ours », confirmation que l'ensemble des rapprochements cités rend du reste superflue.

Or, ce mot indo-européen manque complètement en slave, en baltique et en germanique, et il y est remplacé par des périphrases et des qualificatifs :

v. sl. *medvědi*, serbe *mèdvjed*, slov. *médvęd*, r. *medvéd'*, v. pol.

miedz'wiedz' (Los', *Složnyja slova*, p. 117 ; pol. mod. *niedz'wiedz'*), v. tch. *medvied* (tch. mod. *nedvěd* ; v. Gebauer, *Hist. mluv.*, I, p. 444) ; le mot signifie étymologiquement « mangeur de miel » et répond pour la forme à véd. *madh(u)v-ád-* ; il sert en slave à nommer l'ours mâle. C'est un des très rares composés qui soient vraiment slaves communs, et la composition n'y est plus qu'à peine sensible en slave ; on notera les formations altérées par étymologie populaire, r. dial. *vedmé'd'*, pet. r. *vedmíd'*.

v. sl. *mečĭka*, bulg. *méčka*, s. *mèčka*, r. *méčka* ; ce mot, étranger aux dialectes slaves occidentaux, désigne la femelle, qu'on nomme aussi *medvědica* sur tout le domaine slave; le serbe lui a donné un neutre *mèče* « ourson » et le bulgare un masculin. Le lit. *meszkà* « ours (en général) » doit être un emprunt au russe, avec un léger changement de sens ; il est inutile de recourir, avec M. Brückner (*Lituslav. stud.*, I, 108), à un polonais *mieszek*, qui serait plutôt un emprunt du polonais au lituanien.

lit. *lokys* (accus. *lõkį*; fém. *lókė*), lette *lācis*. Il y faut sans doute voir le « lécheur », cf. lit. *lakù* « je lèche », comme le suggère R. Gauthiot ; le vocalisme long de la voyelle radicale est le même que celui de lit. *žõdis* « parole » en face de *žadù* « je parle », ou de lette *nēsis* « joug » en face de *nest* « porter » ; pour la valeur de la formation, on compare naturellement lit. *gaidys* « coq » (littéralement « chanteur »), etc., v. Leskien, *Bildung der Nomina*, 295 et suiv.

v. pruss. *clokis* Voc. ; le *c* initial n'est pas une faute, car l'article immédiatement suivant du Vocabulaire est : czidelber « *caltestisklokis* », avec cette même gutturale. L'*o* du Vocabulaire répond à lit. *o* (ancien *õ* ou *ā*). Le mot est donc à rapprocher du type lit. *krõkti* « grogner », *klegėti* « rire bruyamment », etc., et signifie « grogneur » ; c'est un des mots de la grande famille des termes à *kr-* ou *kl-* initial, désignant des bruits divers.

v. h. a. *bero*, v. angl. *bera* ont été depuis longtemps rapprochés de lit. *bèras* « brun », etc. ; car on sait que, dans le roman de Renart, l'ours porte en français le nom de *Brun*, en allemand celui de *Braun*, en anglais celui de *Bruin* ; en norrois, le mot ne paraît être représenté que par v. isl. *berserkr*, nom d'une sorte de vête-

ment de guerrier (litt. « vêtement d'ours ») ; le nom norrois de l'ours est v. isl. *bjǫrn*, suéd. *björn*, qui est d'une formation différente ; quant à v. angl. *beorn*, voir Uhlenbeck, P. B. S. B., XXIX, 332 ; Falk et Torp, *Etymolog. ordbok*, p. 58.

Ces dénominations rappellent celles qu'on rencontre chez des populations du Nord de l'Europe, telles que les Estoniens, les Finlandais, les Lapons, qui évitent d'appeler l'ours par son nom et qui le qualifient de « la gloire de la forêt », « le vieux », « la superbe patte de miel », « le poilu », « le pied large », « le mangeur de fourmis blanches », etc. On sait d'ailleurs que, d'une manière générale, l'un des tabous de vocabulaire les plus fréquents porte, durant la saison de chasse, sur le nom de la bête qu'on chasse. Chez les Celtes où le nom de l'« ours » n'a pas disparu comme on vient de le voir, on retrouve des périphrases analogues ; le moyen gallois a *melfochyn*, plur. *melfoch* « ours », litt. « porc à miel », terme qui se retrouve jusque dans un dictionnaire moderne ; l'irlandais a *mathgamain* et simplement *math*, gén. *matho* « ours », à côté de *maith* « bon », gaél. écoss. *math* « bon », comme le montre M. J. Rhys, *Celtae and Galli* (*Proceed. of the Brit. Academy*, II), p. 4.

La glose κυνοῦπες (l. κυνουπεύς ?)· ἄρκτος, Μακεδόνες Hes. indique aussi qu'on évitait en Macédoine le nom propre de l'« ours ». M. O. Hoffmann, *Die Makedonen*, p. 43, y a reconnu un mot apparenté à gr. κνωπ- « animal sauvage ». Les témoignages ne permettent pas de déterminer si cette dénomination a provoqué en Macédoine l'élimination du vieux terme indo-européen, ou si les deux coexistaient comme en celtique.

On est donc conduit à supposer que c'est un « tabou » qui a entraîné la disparition du nom indo-européen de l'ours en slave, en baltique et en germanique ; M. Schrader a déjà émis, sous une forme un peu vague, une idée analogue en s'appuyant sur des faits cités par M. O. Keller (voir son *Reallexikon*, 1re édition, sous le mot *Bär*).

Et en effet on ne voit pas quelle autre cause aurait pu entraîner la disparition du mot.

Il n'y avait pas de cause linguistique : le mot indo-européen

est un thème en *-o-*, donc d'un type courant en indo-européen et conservé partout ; c'est un thème dissyllabique et qui, par suite, n'était ni trop bref ni trop long pour se maintenir.

Il n'y avait pas non plus de cause extérieure : l'animal se rencontrait partout, et sur le domaine des Slaves, des Baltes et des Germains plus que partout ailleurs ; il n'affecte pas de formes très diverses appelant des noms très diversifiés ; et, au surplus, les noms nouveaux ne font pas allusion à des variétés différentes les unes des autres.

On est donc bien en présence d'un « tabou » ; et ce qui achève de le démontrer, c'est le fait, surprenant au premier abord, qu'un autre groupe de langues, le groupe finno-ougrien, n'a aucun mot commun pour désigner l'ours, et que les dénominations de cet animal bien connu de tous les peuples qui parlent ces langues y sont ou empruntées ou périphrastiques et analogues à celles du slave, du baltique et du germanique (d'après une communication de R. Gauthiot à la Société de linguistique de Paris, 24 mars 1906 ; communication suscitée par l'exposé des faits indo-européens qu'on vient de passer en revue).

Étant donné que l'étude du nom de l'« ours » a révélé le rôle des tabous de vocabulaire dans l'histoire des mots indo-européens, il est permis de rechercher si certaines autres particularités ne s'expliqueraient pas de la même manière.

Le serpent est l'un des animaux dont le nom est le plus souvent « tabou ». Or, on constate d'abord que les noms indo-européens du « serpent » n'ont chacun qu'une médiocre extension dialectale et ne se rencontrent que dans un petit nombre de langues géographiquement voisines :

1° skr. *áhiḥ*, zd *aži̯š*, arm. *iž*, gr. ἔφις (ou ἔχις ?), terme oriental et hellénique ;

2° sl. **ǫžь* (r. *už*, pol. *wąż*), lit. *angìs*, lat. *anguis*, terme de la région centrale, inconnu à l'indo-iranien et au grec ;

3° got. *nadrs*, irl. *nathir*, lat. *natrix*, terme purement occidental.

Et, de plus, le serpent est souvent désigné par des épithètes : « rampant » : skr. *sarpáḥ*, lat. *serpens*, gr. ἑρπετόν, alb. *g'arper* ;

ces mots ne se recouvrent pas les uns les autres et n'ont de commun que la racine ; la racine même est différente dans arm. *zerun* à côté de *zeṙal* « ramper » et dans v. h. a. *slango* à côté de *slingan* « ramper ». Il faut évidemment quelque raison particulière pour qu'on désigne le serpent par le « rampant », là où il existe un mot pour dire « serpent ».

« terrestre » : v. sl. *zmĭji* et *zmĭja* (cf. *zemlja* « terre » ; le mot *zmĭjĭ* est dérivé du thème à suffixe zéro **g'hem-* d'où *zemlja* lui-même est tiré) ; cette étymologie, due à M. Hirt et reprise par M. Lidén (*Archiv f. slav. Phil.*, XXVIII, 38), serait difficile à justifier si une interdiction d'employer le nom propre du « serpent » n'avait obligé à recourir à une périphrase obscure en elle-même, mais qui devenait claire par le fait qu'elle était le substitut d'un terme notoirement interdit, au moins dans certaines conditions. Le skr. *uŕagaḥ* litt. « qui va sur le ventre » se justifie peut-être de même.

« vert » : lit. *žaltys*.

« répugnant » : v. sl. *gadŭ* « ἑρπετόν (ce mot désigne tout animal nuisible dans Euch. 59 a) ; cf. pol. *żadny*, etc. ; v. Zubaty, *Archiv. f. slav. Phil.*, XVI, 422, et Brugmann, I.F., V, 375.

On peut même se demander si le skr. *nāgáḥ* « serpent » ne serait pas plutôt l'équivalent phonétique de v. sl. *nagŭ*, lit. *nužas* « nu » qu'un mot apparenté au groupe germanique de v. isl. *snákr* ; si skr. *nāgáḥ* appartient à ce dernier groupe, comme on l'a supposé, il signifierait « rampant » à en juger par v. h. a. *snahhan* « ramper ». Mais, comme skr. *nāgáḥ* désigne aussi un autre animal non velu, l'« éléphant », il ne paraît pas douteux que le sens premier de *nāgáḥ* ne soit « nu », et que *nāgáḥ* ne réponde exactement à sl. *nagŭ*.

Le nom indo-européen de la « souris » est établi par le rapprochement bien connu : skr. *mūṣ-*, pers. *mūš*, v. sl. *myšĭ*, alb. *mī*, gr. μῦς, v. h. a. *mūs*, v. isl. *mūs*, lat. *mūs*. Il ne manque que dans deux langues où il est remplacé par des mots se rapportant à la couleur de l'animal ; en baltique on a lit. *pelẽ*, lette *pele* (et v. pruss. *peles* « muskel » Voc., qui atteste indirectement l'existence de *pelẽ* et montre que ce mot a hérité de tous les emplois de

mūs-, aussi bien du sens de « muscle » que de celui de « souris »), cf. lit. *pelėti* « moisir », *pìlkas* « gris », gr. πολιός, πελιδνός, etc. ; en celtique, on a irl. *luch* (gén. *lochat*), gall. *llyg* et *llygoden*, corn. *logoden*, bret. *lôgôden*, en face de irl. *luch* « noir », gall. *llwg* « pâle » (V. Henry, *Lexique du breton moderne*, sous *lôgôden*, signale en même temps que celle-ci une autre étymologie qui semble moins vraisemblable, surtout si l'on tient compte de lit. *pelė̃*). Or, là où l'on évite de nommer la souris, en Suède par exemple, on l'appelle « la petite grise ».

L'existence des « tabous » pourrait peut-être servir aussi à expliquer les discordances que présentent les noms du « renard » dans les diverses langues indo-européennes (v. Schrader, *Real-lexikon*, sous *Fuchs*).

Dans le recueil *Kulturgeschichtliches aus der Tierwelt*, p. 30, M. O. Keller montre comment les noms du « crapaud » varient d'une langue à l'autre, par ex. gr. φρῦνος, φρύνη « brun », lat. *būfō* (mot dialectal, mais sans doute ancien, v. Niedermann, BB., XXV, 83 et suiv.) et *rubēta* (d'après la couleur), etc.

Il peut arriver que l'interdiction de vocabulaire s'applique à un animal répugnant ; mais il est au moins autant, sinon plus ordinaire, qu'elle s'applique à la bête de chasse ; c'est ainsi que le vieux nom **elen-* du « cerf » (v. sl. *jelenĭ*, etc.) est souvent remplacé par l'épithète « cornu » : lat. *ceruus* et v. isl. *hiǫrtr*, v. angl. *heorot*, v. h. a. *hiruz* rappellent gr. κέρας et ἔλαφος κεραός, « cerf cornu ». Du nom *ragis* de la « corne », le vieux prussien a tiré de même *ragingis* « cerf » à quoi il oppose *glumbe* « biche » (cf. lit. *glumas* « sans cornes »).

Les exemples précédents sont tous tirés de noms d'animaux, parce que ce sont à peu près les seuls où l'action du « tabou » se laisse encore entrevoir. Mais il a dû y avoir beaucoup d'interdictions se rapportant à d'autres notions. Et, à titre de pure hypothèse, on citera encore un cas assez différent de ceux qui ont été mentionnés.

On a observé que, dans une île malaise, près de Sumatra, il est interdit de parler des yeux durant la saison de la chasse. Il est impossible de ne pas songer à cette particularité — aussi à la

superstition du mauvais œil — quand on voit de quelle manière bizarre le vieux nom de l'œil — dont la forme était, il est vrai, très anomale — a été remplacé en irlandais. Au lieu du nom indo-européen de l'« œil », l'irlandais emploie le nom du « soleil », évidemment parce que le soleil était considéré comme l'œil qui voit tout (Avesta, Yasna, I, 11 ; Homère, Γ, 277) : en regard de gall. *heul*, corn. *heuul*, bret. *héol* « soleil » (cf. gr. ἥλιος, got. *sauil*, lit. *sáulė*, etc.), l'irlandais a *súil* « œil » ; une pareille déviation de sens ne devient naturelle et explicable que si l'on admet que le nom propre de l'« œil » a été éliminé, et en effet ce nom ne se retrouve pas plus en brittonique qu'en gaélique.

La forme germanique (got. *augo*, etc.) est trop semblable à lit. *akìs*, v. sl. *oko*, lat. *oculus*, hom. ὄσσε, etc., pour en être séparée, trop différente pour y être ramenée par aucun procédé connu ; n'y aurait-il pas ici une forme voisine de l'ancien nom à laquelle on aurait recouru parce que ce nom lui-même aurait été « tabou » ? alors toute recherche directe de l'étymologie de got. *augo* serait inutile et vaine ; peut-être a-t-on affaire à quelque déformation ou à quelque adaptation d'un mot de sens plus ou moins éloigné. On remarquera d'ailleurs que, tout en appartenant pour la plupart à une même racine, les noms de « l'œil » diffèrent d'une langue indo-européenne à l'autre.

Il est curieux que le duel zend *aši*, qui répond exactement à skr. *akṣí*, ait été attribué au vocabulaire ahrimanien, sans doute dès les gâthâs (Yasna, XXXII, 10) ; pour les êtres bons, on lit dans l'Avesta *čašma*, qui s'est conservé jusqu'à l'époque actuelle (pers. *čašm*), ou *dōiθrəm*, formation nouvelle et peut-être artificielle, qu'ignorent les dialectes modernes ; à qui douterait que *čašm* désigne le « bon œil » par opposition au mauvais œil, il y a un fait décisif à opposer : en vieux perse, le nom de l'« œil » est (*h*)*učašma*, c'est-à-dire, comme l'a bien vu M. Weisbach, le « bon œil ». De même le vieux mot zend *uši* « les (deux) oreilles » (cf. v. sl. *uši*) est ahrimanien, ainsi que *karəna-* = skr. *kárṇa-* « oreille », le mot employé pour les êtres bons est le terme exclusivement propre à l'iranien zd *gaošō*, v. pers. *gauša*, qui se retrouve encore aujourd'hui dans persan *gōš*. Par ailleurs, au

contraire, c'est le terme indo-iranien qui sert pour les êtres bons, ainsi zd *zastō* « main » = skr. *hástaḥ*, ou zd *pad-* = skr. *pad-* « pied », et ces mots survivent encore maintenant, ainsi persan *dast* (v. perse *dasta*) et *pāi*; les termes ahrimaniens sont alors empruntés à un mot ancien, mais accessoire, comme zd *gava-* « main » (v. Lidén, *Arm. stud.*, p. 120), ou artificiellement fabriqués pour les besoins d'un passage particulier, comme les ἅπαξ zd *dvarəθra-* et *zbaraθa-* « pied »; le vocabulaire ahrimanien de l'Avesta, n'a, comme on le voit, aucune unité; si on l'étudiait en détail, il faudrait ajouter aux types cités ici celui des cas où le terme propre est employé seulement pour les êtres ahrimaniens, et est remplacé pour les êtres bons par un euphémisme : les êtres ahrimaniens « meurent » (racine *mar-*), les êtres bons « passent ».

Il est donc nécessaire de supposer que certaines interdictions ont porté sur le nom de l'œil et aussi sur celui de l'oreille **aus-*, **us-*, qui se trouve n'être pas représenté en sanskrit, qui est ahrimanien dans l'Avesta et qui a disparu de bonne heure dans les dialectes iraniens.

L'expression de l'idée de « droit » se fait dans presque tous les dialectes indo-européens au moyen de diverses formations d'un même élément radical **deks-* qui se rencontre depuis l'indo-iranien jusqu'à l'italo-celtique, en passant par le baltique, le slave, l'albanais, le germanique et le grec; l'arm. *aǰ* « droit » diverge seul (v. Lidén, *Arm. stud.*, p. 75 et suiv.). Au contraire, pour l'idée de « gauche », il y a plusieurs expressions distinctes, dont chacune n'a qu'une faible extension :

1° skr. *savyáḥ*, zd *haoya-*, v. sl. *šujĭ*;

2° v. sl. *lěvŭ*, gr. λαιός; lat. *laeuos*;

3° gr. σκαιός, lat. *scaeuos*; cf. peut-être lit. *kairė̃*.

Et non seulement on a ainsi trois termes indo-européens distincts au lieu d'un qu'on a pour l'idée de « droit », mais surtout chaque langue recourt à des artifices pour exprimer cette notion qu'on préférait ne pas nommer directement; le grec a εὐώνυμος et ἀριστερός, l'Avesta *vairyāstara-* (v. Bartholomae, *Altiran. Wört.*, sous ce mot), etc.; on trouvera les principaux moyens dont on s'est servi chez Schrader, *Reallexikon*, sous *Rechts und Links*.

Si les noms de maladies et d'infirmités, même des plus fréquentes et des plus connues, comme la boiterie, la cécité, la surdité, diffèrent d'une langue à l'autre et ne sont réductibles que rarement à des formes indo-européennes, c'est évidemment qu'on évitait ces noms.

D'une manière générale, l'absence d'un nom indo-européen commun dans des conditions où a priori on s'attendrait à en trouver un appelle toujours une explication, et ce n'est pas forcer l'importance du principe des interdictions linguistiques que d'attribuer à des sortes de « tabous » l'inexistence d'un terme indo-européen pour une notion qui en devrait normalement avoir un. Mais, comme ces interdictions ne sont pas directement attestées, on doit aussi se garder d'exagérer le rôle de pareilles explications et même de rien affirmer à cet égard d'une manière absolue[1].

1. Il importe de compléter cet article par celui que Gauthiot a consacré aux noms de l'abeille et de la ruche, *Mémoires de la Société de linguistique*, XVI, p. 264 et suiv., et par celui que M. Vendryes a consacré à latin *mundus*, *ibid.*, XVIII, p. 305 et suiv.

A PROPOS
D'UN
RÉCENT DICTIONNAIRE ÉTYMOLOGIQUE
DU FRANÇAIS [1]

Faire l'étymologie d'un vocabulaire donné, c'est faire l'histoire de ce vocabulaire entre deux dates. Ici comme en toute histoire, il dépend de l'historien de choisir les dates entre lesquelles il juge bon de suivre le développement. Mais il faut choisir. En matière de français, on a pratiquement le choix entre deux dates initiales : on peut partir du latin de l'époque impériale, ou de l'indo-européen. On s'accorde habituellement pour partir du latin d'époque impériale, parce que les romanistes qui s'occupent du français ne sont pas d'ordinaire assez comparatistes pour faire l'histoire des mots entre l'indo-européen et le latin historique et parce qu'en partant de l'indo-européen, on aurait une histoire trop longue et trop complexe.

La question qu'on pose vulgairement aux linguistes : « Quelle est l'origine d'un mot ? », n'a pas de sens précis.

Ce qui est essentiel dans un dictionnaire étymologique, c'est de déterminer les voies qu'ont suivies les mots. On procède souvent comme si les changements de sens s'expliquaient par des faits psychiques. En réalité, les variations de sens sont déterminées par des circonstances de fait, de caractères infiniment divers. Pour faire l'étymologie d'un mot, il faut donc avant tout savoir comment il a pénétré en français et dans quels milieux il a été employé.

Ce qui importe pour expliquer *escobar*, *escobarderie*, ce n'est pas

1. Extrait du *Bulletin de la Société de linguistique*, XXI, p. 81 et suiv. (année 1918).

qu'Escobar ait été un casuiste espagnol du XVII^e siècle, c'est que Pascal en ait fait l'une des figures des *Provinciales* : ce n'est pas le traité de casuistique d'Escobar qui rend compte du sens pris par le mot en français, c'est le parti qu'en a tiré Pascal et le succès des *Provinciales*.

Dire que « *cipaye* et *spahi* sont deux formes différentes d'un même mot persan », c'est enseigner une simple curiosité. On ne dira quelque chose d'utile que si l'on indique par suite de quels événements l'un vient de l'Inde comme l'anglais *sepoy*, et l'autre de la Turquie et de l'Algérie.

Il est vain de montrer des dérivations sans indiquer où, quand et comment elles se sont faites, et des changements de sens sans en marquer les conditions historiques. A quoi bon ramener à un primitif unique le plus de mots qu'il est possible, si l'on ne signale pas le passé propre de chacun de ces mots ? La plupart du temps, il s'agit de faits très divers, et l'unité d'origine mise en évidence est de beaucoup ce qu'il y a de moins intéressant pour expliquer la forme et le sens des mots considérés.

Soit, par exemple, l'article *seul* du dictionnaire. L'auteur, après avoir signalé que fr. *seul* est le latin *sōlum*, ajoute : « Dérivés savants : **solitude**, **solitaire**, **soliloque** (v. *locution*) ; **solo**, mot italien, **soliste** ; **désoler** ; . . . ».

En ce qui concerne *solitude* et *solitaire*, l'expression « dérivés savants » ne donne pas une idée claire de la réalité : ces mots français sont les adaptations normales de mots latins, dérivés de *sōlus* : *sōlitūdō* et *sōlitārius* ; il s'agit d'un caractère habituel du français littéraire, et la première chose à faire dans un dictionnaire étymologique du français serait de noter d'une manière constante ce procédé ordinaire qui consiste en ce que les dérivés français sont, en très grande partie, tirés non de mots français, mais de mots du latin écrit.

Les Latins sentaient sans doute encore le lien entre *sōlus* et *dēsōlāre* ; mais il va de soi que, en faisant l'emprunt, le français a radicalement isolé *dēsōlāre* de *seul*, et même de *sōlus* ; l'évolution du sens le montre bien. On ne saurait mettre sur une même ligne *solitude*, *solitaire* et *désoler*.

L'italien *solo* n'est pas un dérivé ; traiter de même l'emprunt à la langue musicale italienne *solo* et des emprunts au latin écrit comme *solitude* et *solitaire*, c'est dissimuler deux faits historiques de caractère distinct.

L'auteur du dictionnaire ajoute : « A la forme archaïque *sollum* et au sens également archaïque de « tout, chaque », voisin du sens de « seul », se rattachent : latin *solennem*, d'où **solennel**... ; **solliciter** (et les mots apparentés, v. *citer*) ; l'adjectif *solidum*... » Le lecteur non averti conclura de là sans doute que *sōlus* a eu autrefois le sens de « tout, chaque », ce qui est faux ; et il croira que le rapprochement de *sōlus* et de *sollus* est chose sûre, universellement admise ; or, ce rapprochement n'est ni évident, ni communément admis ; le seul mot latin dont la parenté avec *sollus* « entier » (et non pas « chaque » : le fr. *tout* prête à une ambiguïté de sens) soit certaine est *saluos*, qui se retrouve dans fr. *sauf*, mais que l'auteur ne mentionne pas ici. Une étymologie aussi douteuse — presque sûrement fausse — que celle de *sōlus* rapproché de *sollus* ne doit pas figurer dans un livre de vulgarisation.

Au point de vue latin, les trois mots, *solennis*, *sollicitus* et *solidus* n'avaient déjà plus aucun rapport entre eux ; il y a donc grand inconvénient à les rapprocher dans un dictionnaire étymologique du français.

L'histoire de *solidus* est curieuse ; mais en la présentant sous un article *solidus*, subdivision d'un article *sollus*, rattaché lui même à un article *sōlus*, on la fausse entièrement. En effet, il y a les éléments suivants, rigoureusement indépendants les uns des autres :

1° Un adjectif emprunté au XVI^e^ siècle au latin écrit, *solide*, avec ses dérivés.

2° Le mot *solidaire* pris à la langue juridique au XVI^e^ siècle : il est curieux que ce mot n'existe pas en latin ancien.

3° Le lat. *soldus*, doublet latin de *solidus*, au sens de « pièce de monnaie ». La traduction de *solidus* par « monnaie en métal solide » n'a pas de sens saisissable ; en réalité, on sait que *solidus*,

appliqué à un métal, signifiait « massif », et qu'on s'est servi de *solidus* à l'époque impériale pour désigner une monnaie d'or. La déchéance de *soldus* qui, sous l'Empire, désignait une pièce d'or et qui, dans les langues romanes, est devenu le nom du billon le plus humble, le *sou* vaut d'être signalée ; il y a là un développement remarquable.

4° A cette valeur de *sol(i)dus*, se rattache le sens de « solde, prix payé pour un service » ; l'italien *soldo* signifie à la fois « sou » et « solde ». Et c'est à l'italien *soldo* qu'est pris le mot français *solde*, avec un genre différent. Mais, de ce sens de *sol(i)dus*, a été tiré un verbe *solidāre* « donner en solde, payer entièrement », d'où *solidáta*, prov. esp. *soldada*, v. fr. *soudée ;* il y aurait à chercher en quelle mesure le genre de fr. *solde* serait influencé par de vieux mots français ; et, si l'influence est admise, le fr. *solde* apparaîtrait comme une contamination d'anciens mots français et d'un mot italien. De *soudée*, vient un français *soudoyer*, tandis que *soudard* est dérivé du v. fr. *soude*. — Il faudrait aussi examiner en quelle mesure *solder* au sens de « apurer » un compte, et, par extension, « achever l'écoulement d'une marchandise », est un emprunt à une forme latine *sol(i)dāre*, et en quelle mesure l'italien *soldare* a pu agir : on sait que les termes français de banque sont d'origine italienne. — La pire erreur qu'on puisse commettre en exposant des étymologies, c'est de laisser croire au lecteur que l'histoire des mots est chose simple. Les influences les plus diverses s'y croisent.

5° De *soldare*, l'italien a fait *soldato*, qui désignait le mercenaire des armées italiennes. Le français a emprunté le mot au XVI^e^ siècle ; mais, comme les armées avaient en France un tout autre caractère qu'en Italie, et que, d'ailleurs, la formation n'était pas intelligible en français, le mot a pris immédiatement une noblesse qu'il n'avait pas en italien. — Le mot *soldatesque*, demeuré plus italien, a conservé une nuance de mépris (*soldatesque* ne vient pas de fr. *soldat* ; il est italien).

6° Le lat. *sol(i)dāre* a abouti à it. *sodare*, fr. *souder* ; en français, *souder* est devenu un terme de métier, purement technique.

Mettre dans un même article *solide*, *solidaire*, *sou*, *solde* (et

solder), *soldat*, *souder*, c'est réunir six — et même sept — mots distincts, et dont l'histoire est distincte. Constater que ces six ou sept mots différents du français se rattachent à un même original latin, c'est se borner à une curiosité d'intérêt médiocre, de portée presque nulle. Ne rien dire de plus, ou presque, c'est éliminer tout ce qui, dans l'étymologie, a un intérêt.

Pour faire l'histoire des mots, il faut poser le point de départ roman commun ; or, on ne peut le faire que par la comparaison des langues romanes. En face des mots français, il faut donc mettre les correspondants des autres langues néo-latines. Ce n'est pas *balneum* qui rend compte du fr. *bain ;* c'est une forme **baneum* que supposent fr. *bain*, prov. *banh*, it. *bagno*, esp. *baño*, port. *banho*, et qui se retrouve dans l'emprunt slave *banja* (passé au féminin). Ces rapprochements ont du reste une haute valeur pratique : ils montrent le parallélisme des langues romanes et en font ressortir l'unité profonde.

Seule, la comparaison des langues romanes fait comprendre d'où est parti le français. Toutes les langues romanes autres que le roumain ont éliminé *cerebrum*, et ont remplacé *cerebrum* par le diminutif *cerebellum* ; on attribue ce mot au latin populaire ; en réalité, *cerebellum* désigne la « cervelle » en tant qu'elle se mange ; comme le nom roman du « foie », ç'a été d'abord un terme de cuisine et de gastronomie ; le rapprochement valait d'être fait. Quant à l'opposition de *cerebellum* et de la forme réellement vulgaire **cerebella*, elle n'est établie que pour le gallo-roman : hors du français et du provençal, un représentant de *cerebella* n'est signalé que dans l'Italie du Nord.

Faire un véritable dictionnaire étymologique du français est, actuellement, impossible ; car on est loin de connaître l'histoire des mots français depuis l'indo-européen ou simplement depuis l'époque de l'Empire romain jusqu'à présent. Mais les faits actuellement connus permettraient d'en tracer l'esquisse, de faire entrevoir les questions, et, par là même, de pousser à faire les travaux préparatoires d'une manière rapide et systématique. Il y a là un véritable devoir pour les peuples de langue romane en général, pour les Français en particulier.

A PROPOS
DES
NOMS DU VIN ET DE L'HUILE[1]

Entre la morphologie et le vocabulaire il y a une grande différence : la morphologie est un système qui n'admet que très malaisément l'introduction d'éléments étrangers, non livrés par la tradition de la langue même ; le vocabulaire se compose de mots dont sans doute la valeur ne se laisse définir précisément que par rapport à d'autres mots, mais qui néanmoins sont indépendants les uns des autres, et, par suite, il peut accueillir en quantité illimitée des éléments ne provenant pas de la tradition propre de la langue. Dans toute langue, la morphologie est la part traditionnelle ; le vocabulaire réfléchit les diverses influences de civilisation. L'histoire du français en fournit un exemple saisissant.

Toute la grammaire du français se compose de procédés dont les uns continuent simplement des procédés latins — c'est la moindre part — et dont les autres — les plus nombreux — sont obtenus à l'aide d'éléments latins, ou, du moins, développés en français même et par l'effet de conditions existant dans la langue. L'opposition entre *chante* et *chantez* réfléchit celle entre *canta* et *cantate* ; l'opposition entre *neuf* et *neuve* celle entre *nouum* et *nouam*. Des formes nouvelles telles que *je chanterai* ou *j'ai chanté* se ramènent à des groupes tels que *ego cantare habeo* ou *ego habeo*

1. Extrait de *The French Quaterly*, II (1920), p. 1 et suiv. Le fond des idées exposées ici a fait déjà l'objet d'une note de caractère technique dans les *Mémoires de la Société de Linguistique*, XV, p. 161 et suiv. Depuis, des hypothèses analogues ont été proposées par divers savants, notamment par MM. Cuny et Theander.

cantatum qui n'avaient pas en latin le caractère de procédés grammaticaux, mais où l'on trouve l'amorce de procédés devenus grammaticaux en français. Le rôle de la préposition *de* en français est chose nouvelle ; mais *dē* est une préposition latine. La signification grammaticale de l'ordre des mots ne remonte pas au latin ; c'est au cours du développement que s'est créée cette signification, sans qu'une influence extérieure soit intervenue.

Il en est autrement du vocabulaire. Il y a, pour les notions les plus usuelles, un fonds principal qui continue le fonds latin : des mots accessoires tels que les pronoms, les noms de nombre, les noms d'animaux les plus répandus, les noms de parenté, les verbes indiquant les actes ordinaires comme *boire* ou *dormir*, beaucoup d'adjectifs, etc... Mais, à côté, il y a les mots qui proviennent d'origines diverses. Il a subsisté quelques mots gaulois comme *benne*. Il s'est introduit beaucoup de mots germaniques comme *guerre*, *choisir*, *garder*, des mots arabes ou américains, venus pour la plupart par l'intermédiaire de l'espagnol, comme *alcôve*, *chocolat*, etc... Chacune des civilisations avec lesquelles les Français se sont trouvés en contact, médiat ou immédiat, a ainsi laissé dans le vocabulaire français quelques traces de son influence.

Pas plus dans le langage que dans la technique, les civilisations abolies ne meurent tout entières. La civilisation, relativement avancée, des Gaulois a laissé chez ses voisins bien des traces. La grande métallurgie du fer qui caractérise l'époque gauloise a fourni aux langues germaniques le nom du « fer » **isarnan* qui est manifestement celtique. Les Gaulois avaient largement développé l'usage des chars, hérité de l'époque indo-européenne ; les Romains qui les ont refoulés, puis conquis, et qui leur ont transmis l'ensemble de leur civilisation, ont accepté la terminologie gauloise relative aux chars : le gaulois *carrus* est devenu un mot latin courant, qu'ont conservé les langues romanes et que les légions romaines ont porté au loin, jusqu'en Arménie où on le retrouve intact.

*
* *

Avant l'arrivée des tribus de langue indo-européenne qui ont apporté en Grèce les parlers appelés à devenir la langue grecque, il y avait dans le bassin de la mer Egée une civilisation brillante dont de récentes trouvailles faites surtout en Crète et qui ont rendu célèbre le nom de M. Evans, ont révélé des restes nombreux et caractéristiques. Au milieu de cette civilisation élégante et riche, les « Hellènes » — pour leur donner le nom sous lequel on connaît leurs descendants — sont apparus comme des barbares venus du Nord ; leur arrivée a été suivie d'une sorte de « moyen âge » relativement rude. Mais de même que les Germains ont beaucoup pris à la civilisation gréco-romaine que leur venue a contribué à disloquer, les Hellènes n'ont pas manqué de beaucoup emprunter à la langue — ou aux langues — par laquelle s'exprimait la civilisation « égéenne » au second millénaire avant l'ère chrétienne.

Ces langues sont inconnues. Il est permis de supposer qu'elles appartenaient au groupe d'Asie Mineure : lydien, lycien, carien, dont on n'a que peu de monuments, — et ce peu n'est interprété que d'une manière partielle, incertaine, — si bien qu'on ne saurait faire aucune théorie des anciennes langues de l'Asie antérieure. Les fouilles de Crète ont apporté de vieilles inscriptions qui sans doute sont rédigées dans la langue des Crétois préhelléniques, de ces gens qu'Homère qualifie de vrais Crétois (*Eteokrêtes*) ; mais elles ne sont pas déchiffrées ; si elles l'étaient, on ne les comprendrait sans doute pas mieux qu'on ne comprend les inscriptions en caractères grecs, mais en une langue inconnue, qu'on a trouvées en Crète ou que les textes non grecs écrits en alphabet cypriote.

En somme, on ne sait rien ou presque rien des langues qui ont servi d'organe à la civilisation égéenne du second millénaire avant l'ère chrétienne.

*
* *

Une seule chose est sûre : les tribus de langue indo-européenne qui ont envahi l'extrémité de la péninsule balkanique et tout le

bassin de la mer Egée ont beaucoup pris à la vieille culture égéenne. Avec les choses, ils ont dû prendre beaucoup de mots.

Et, en effet, si l'on examine le vocabulaire grec, on constate que beaucoup des mots, et en particulier des substantifs, qu'il comprend ne s'expliquent pas au moyen de ce que l'on sait du vocabulaire indo-européen, ou même ont un aspect qui les dénonce comme n'étant sans doute pas indo-européens.

Un mot comme *labyrinthos*, dont il n'est question qu'à propos de l'ancienne Crète ou de l'Égypte, n'a rien d'indo-européen. L'élément *-intho-*, qui paraît avoir servi à la formation, se retrouve dans des noms propres tels que *Korinthos*. Or, les noms propres grecs de lieux ne s'expliquent en général pas par le grec, c'est-à-dire qu'ils appartiennent à des langues parlées en Grèce avant l'arrivée des nouveaux conquérants.

Ceci fait apparaître qu'un substantif comme le nom *asamin-thos* de la « baignoire » qu'on lit déjà chez Homère, et qui ne s'explique pas par l'indo-européen, est d'origine « égéenne ». Dès lors il est permis de se demander si le nom *balaneion* du « bain » que la forme ne dénonce pas pour étranger, mais pour lequel on n'a pas trouvé d'étymologie indo-européenne satisfaisante, ne serait pas aussi « égéen ». Comme ce mot a passé en latin, *bali-neum*, *balneum*, et de là aux langues romanes (français *bain*, etc.) et avec sa forme latine vulgaire, au slave (*banja*), le mot « égéen » est devenu un mot presque européen.

Le « cyprès » est un arbre de la région méditerranéenne. Son nom grec *kyparissos* n'a rien d'indo-européen, ni par la racine, ni par l'élément de formation *-isso-* qui, au contraire, est courant dans le monde égéen. Le grec doit donc le mot à une langue « égéenne » antérieure. Et la forme *cupressus* du mot latin correspondant, qui s'explique mal en partant de la forme grecque, provient sans doute d'un emprunt fait par le latin, non au grec, mais, directement ou indirectement, à la langue à laquelle a puisé le grec lui-même.

Le nom grec *thriambos* désigne un cortège bachique et le nom latin *triumphus* (que les Grecs ont rendu par *thriambos*) ont de même été reconnus pour probablement « égéens » et ceci con-

duit à envisager la possibilité d'une origine « égéenne » pour des mots savants tels que ceux de l'*ïambe* et du *dithyrambe*. Il est aisé de trouver des exemples plausibles, et même probables, de mots grecs importants ayant une origine égéenne. La démonstration complète échappe le plus souvent.

*
* *

La nation qui parlait l'indo-européen habitait sans doute une région assez septentrionale, où prospéraient le hêtre et le bouleau, mais où l'on ne cultivait pas la vigne. La seule boisson fermentée dont les restes conservés de la langue indiquent l'existence est l'hydromel, dont l'anglais *meat* et l'allemand *met* ont gardé jusqu'à présent le nom ancien, comme aussi les langues slaves (russe *mëd*, etc.). Ceux des peuples de langue indo-européenne qui ont appris à connaître le vin se sont parfois servis du vieux nom de l'hydromel pour désigner le vin ; c'est ainsi que le grec *méthy* désigne le « vin », et que les mots qui indiquent l'ivresse s'y rattachent : *methê* « ivresse », *methyô* « je suis ivre », *methyskomai* « je m'enivre » ; ceux-ci n'ont pas eu à changer de sens, avec le produit qui donnait l'ivresse.

Mais, à côté de ce mot qu'il adaptait pour le sens, le grec en a emprunté un autre, presque évidemment « égéen », à savoir *woinos* (la forme *oînos* de l'ionien-attique s'explique par la chute de *w* initial qui est de règle dans ce dialecte). Ce mot n'est pas isolé : on en retrouve une forme apparentée dans les langues italiques : *uīnum* en latin, *vinu* en ombrien : le genre neutre du mot et l'*ī* de la première syllabe indiquent assez que le mot italique n'est pas emprunté au grec, comme on l'a trop souvent répété. L'arménien, qui n'a emprunté directement des mots grecs que depuis l'époque chrétienne, a un mot *gini*, qui repose sur une ancienne forme analogue à *woînos* du grec et qui doit, par quelque intermédiaire inconnu, se rattacher au même original égéen ; de l'arménien, ce mot a passé au géorgien, à une date antérieure à la période historique de l'arménien. Le même nom du vin se retrouve enfin dans tout l'ensemble du groupe sémitique (arabe

waynun, etc), sans qu'il y ait lieu d'examiner ici si l' « égéen » a emprunté le mot au sémitique, ou inversement. Ainsi un nom méditerranéen du « vin » figure dans la plupart des langues qui touchent par quelque côté à la Méditerranée ; et, sous la forme latine, ce mot s'est imposé à l'Europe entière, c'est-à-dire aux langues celtiques, germaniques et slaves. Si l'on ne posait pas un original égéen, on n'arriverait pas à comprendre d'où le nom de « vin » a pu rayonner ainsi sur l'Europe et l'Asie.

La grande fortune du mot *uīnum* en latin se marque au nombre des mots que le français en a tirés. Le français n'a pas seulement *vin* qui continue *uīnum*. Il a aussi *vigne* qui continue le dérivé *uīnea* de *uīnum*, *vendange*, qui continue un composé *uindēmia*, dont le premier terme est *uinum*, et enfin un ancien juxtaposé *vinaigre*, qui s'explique immédiatement en français, bien que, actuellement, les sujets parlants n'y sentent plus le mot *vin*, et que, d'ailleurs, le « vinaigre » employé ne soit le plus souvent pas du « vin » ayant subi la fermentation acétique. Il y a donc en français quatre mots, actuellement indépendants les uns des autres, qui se rattachent à *uīnum*.

Il est frappant que les noms grecs de la « vigne », *ampelos*, et du « raisin », *botrys*, n'aient pas d'étymologie indo-européenne connue. Ceci vient confirmer l'hypothèse de l'origine « égéenne » du grec *woinos*, et, par suite, du latin *uīnum*.

*
* *

Plus évidemment encore que la vigne, l'olivier est une plante méditerranéenne. Les graisses dont l'indo-européen a le nom sont des graisses animales. Les Grecs n'ont appris à connaître les graisses végétales qu'en entrant en contact avec le monde égéen, où la culture de l'olivier et l'emploi de l'huile d'olive avaient une grande importance.

Le nom que le grec a emprunté a été adapté à des usages grecs : on a fait une forme féminine *elaiwâ* (en attique *elaa*) pour désigner l'arbre, « l'olivier », et son fruit « l'olive », une forme neutre *elaiwon* (en attique *elaion*) pour désigner le produit,

l' « huile ». Mais le mot est d'origine étrangère. Du reste, on en retrouve l'élément essentiel *el-* en arménien (*ewł* enarménien classique), *et* dans les parlers populaires qui ont fourni l'arménien moderne, sous une forme qui ne peut être d'origine grecque, et qui, comme le nom arménien du « vin » remonte, directement ou indirectement, à la civilisation « égéenne ». M. Marcel Cohen me signale même que l'on retrouve en sémitique et en hamitique des formes assez semblables pour désigner des arbres non cultivés. Il y a là toute une recherche curieuse à poursuivre.

L'opposition entre le nom *oliua* de l' « olivier » et de l' « olive » et le nom *oleum* de l' « huile » montre que ces mots latins sont empruntés au grec à la différence de ce que l'on a observé pour le nom du « vin ». La différence de forme entre les mots latins, *oliua*, *oleum*, et les mots grecs dont ils sont issus et les différences entre l'aspect de *oliua* et celui de *oleum* s'expliquent par des particularités connues de la phonétique latine et ne font aucune difficulté. Mais cet accident phonétique a séparé *oliua* de *oleum* pour le sentiment linguistique d'un Romain. Dans une langue comme le français, où les noms de l' « olivier » et de l' « olive » ne figurent que par emprunt à des parlers des régions possédant l'olivier, et où l'huile indigène se fabrique avec d'autres produits végétaux que l'olive, le nom de l'huile a perdu toute attache avec l'olivier Il en est naturellement de même du mot anglais *oil*, qui est d'origine française, ou de l'allemand *öl*. Et quand, comme il arrive souvent maintenant, on se sert du mot *huile* (ou de l'anglais *oil*) pour désigner des huiles minérales, on est bien loin du monde égéen, d'où vient le mot.

L'histoire du nom de l'huile dans les langues romanes est d'ailleurs compliquée. Nulle part, le mot n'a la forme qu'on attendrait s'il s'agissait d'un mot traditionnel. Des influences savantes sont intervenues partout.

*
* *

Il y a sans doute bien d'autres noms encore que ceux de « vin » et de « l'huile » que la civilisation moderne doit à la vieille civilisation égéenne.

Par exemple, il est probable que le nom de la « rose » qui se trouve et en grec et en iranien, et qui, de l'iranien, paraît avoir passé à l'arménien, est « égéen ». En grec, le mot n'a pas été emprunté à la date la plus ancienne : il est curieux que la forme qu'il présente ait un caractère dialectal particulier, celui de l'éolien, et que, par suite, il doive passer pour emprunté à l'éolien dans les autres dialectes. La forme latine *rosa* présente à l'explication de grandes difficultés qui ne se laissent pas résoudre en supposant un point de départ grec du mot latin.

Les langues indo-européennes de l'ancien groupe du Nord-Ouest, du slave et du baltique au celtique et même à l'italique, ont eu un nom de la « pomme » dont l'anglais *apple* par exemple est encore un représentant. Mais les populations qui sont arrivées au contact de la culture égéenne ont sans doute trouvé une pomme plus perfectionnée, dont le nom, *mâlon* (en ionien-attique *mêlon*), apparaît en grec ; le latin l'a emprunté au grec, d'abord sous la forme de quelque parler dorien où *ā* subsistait, *mālum*, puis sous la forme empruntée à l'attique *mēlum*, qui a survécu notamment dans l'italien *melo*. Parmi les parlers italiques, le vieux nom de la « pomme » ne subsiste que dans un nom propre, celui de la ville d'Abella, célèbre par ses pommes, *malifera Abella*, comme dit Virgile, rapprochant ainsi de manière curieuse les deux noms de la « pomme », l'ancien, indo-européen du Nord-Ouest, et le nouveau, emprunté au monde « égéen ».

On a trop peu de données pour déterminer, même approximativement, ce que le grec et le latin, et par là les langues modernes, doivent à l'ancienne culture « égéenne » du second millénaire av. J.-C. Mais le peu que l'on entrevoit montre que cet apport est important.

Ainsi se maintiennent dans les langues les restes de civilisations depuis longtemps abolies et oubliées.

J. GILLIÉRON

ET

L'INFLUENCE DE L'ÉTUDE DES PARLERS LOCAUX SUR LE DÉVELOPPEMENT DU ROMANISME.

M. J. Gilliéron, qui depuis plus de trente ans dirige à l'École des Hautes Études les travaux sur les parlers gallo-romans, n'a rien fait pour se faire connaître du grand public. Son dernier livre, la *Généalogie des noms qui ont désigné l'abeille*, n'est destiné qu'aux spécialistes, et, même pour les gens du métier, si la lecture en est étrangement savoureuse, elle n'est pas facile. Mais il n'y a pas d'homme qui, depuis qu'il enseigne, ait eu plus d'action sur tous ceux qui ont étudié l'histoire des parlers gallo-romans, l'histoire des langues romanes, et, finalement, toute la linguistique historique en général.

Depuis que s'est brisée, à partir du III^e^ siècle après Jésus-Christ, l'unité linguistique latine, chaque province de la France actuelle, chaque localité même a eu son développement linguistique propre. A n'envisager que le parler local, il y a dans la France d'aujourd'hui autant d'objets d'observation distincts qu'il y a de villages. Actuellement, ces parlers locaux, autrefois bien distincts, tendent à s'éliminer très vite. Tandis que, dans la France méridionale, où les parlers diffèrent entièrement du français, ils se maintiennent encore, ils subissent de plus en plus dans toute la France centrale l'action du français commun, dans un rayon de deux à trois cents kilomètres autour de Paris. Mais partout il subsiste des restes notables de l'ancien parler local qui, il n'y a pas plus d'un siècle, était encore employé par presque tout le monde dans chaque village et même dans chaque petite ville.

Il y a d'autant plus d'intérêt à observer ces parlers locaux que,

entre l'ancienne unité latine, qui s'est constituée du temps de l'Empire romain, et l'unité française, qui se constitue maintenant grâce à l'extension du français commun, il n'y a eu en France aucune période d'unité linguistique. En une certaine mesure — qu'il ne faut naturellement pas exagérer — chaque village de France a eu, depuis l'époque latine, son développement linguistique propre. L'évolution du latin dans les divers villages du domaine roman offre assurément le plus bel ensemble qui s'offre à qui veut étudier comment les langues se développent au cours du temps.

Chargé de diriger, à l'École des Hautes Études, les travaux relatifs aux parlers gallo-romans, M. Gilliéron a compris que l'on n'arriverait à rien de définitif en examinant, çà et là, au hasard, tel ou tel parler et en essayant de voir comment il se comporte par rapport au latin : la distance du latin à chaque parler est trop grande, les influences subies ont été trop diverses. On est obligé de remplir par des hypothèses invérifiables un développement trop complexe ; la part faite à l' « interpolation » est trop grande. Et il y a trop de faits communs qui ne ressortent pas dans une étude portant sur une seule localité ou sur une seule région.

M. Gilliéron a donc organisé une enquête systématique portant sur la France entière. Les sommes mises à sa disposition à la fois par une modeste subvention d'État, par un petit prélèvement sur la modique somme que la Ville de Paris met à la disposition de l'École des Hautes Études et par la maison d'édition Champion étaient trop exiguës pour permettre une enquête complète qui, de plus, aurait demandé un temps immense et dont l'examen aurait beaucoup dépassé les forces de l'homme le plus laborieux. M. Gilliéron a choisi un certain nombre de points répartis sur la France entière de manière à donner une idée de chaque district ; un enquêteur, doué d'une très bonne oreille, M. Edmont, a sur chacun des points, fait traduire dans le parler local toutes les phrases d'un questionnaire établi pour permettre de résoudre la plupart des problèmes que l'on envisageait au moment où a été fait ce questionnaire. Une fois toutes

les réponses obtenues, elles ont été reportées sur des cartes, dont chacune indique la forme d'un mot donné sur le millier de points qui ont été enquêtés. Ces cartes sont maintenant publiées, par les soins de la librairie Champion, dont elles sont l'honneur ; elles constituent l'*Atlas linguistique de la France*, qui est un monument unique. C'est la première fois qu'un grand domaine linguistique est décrit dans toute son étendue, et qu'on peut, d'un coup d'œil, apercevoir la façon dont un élément linguistique ancien a été traité, indépendamment, sur des centaines de points différents. Grâce à l'Atlas, toutes les données recueillies soit auparavant soit depuis viennent prendre place dans un ensemble.

Plusieurs des jeunes romanistes qui ont reçu à l'École des Hautes Études l'enseignement de M. Gilliéron ou qui ont subi l'influence de ses publications, libres de tous dogmes, inspirées seulement de l'observation des faits, ont poussé plus avant : ils ont étudié de plus près des domaines restreints : M. Millardet, dans les Landes, M. Bruneau, dans les Ardennes, M. Terracher, dans la Charente, M. Oscar Bloch, dans les Vosges, ont décrit plus ou moins complètement de petites régions, village par village. Les atlas linguistiques de M. Millardet et de M. Oscar Bloch ont montré ce que peut être une étude géographique de parlers poussée jusqu'au bout.

Ainsi a été réalisé un progrès décisif pour la linguistique. Toute linguistique historique est comparative, et seule la comparaison des divers parlers issus d'une même langue commune peut servir à déterminer cette langue commune. Sans doute les langues romanes continuent le latin ; mais la forme du latin qu'elles continuent n'est pas celle que fournissent les textes : qui chercherait à expliquer le nom roman du « foie » par le nom latin *iecur* ou le nom de l'oreille par le nom latin *auris* demeurerait bien embarrassé. Des textes latins fournissent *auricula* ; mais rien, à considérer ces textes, ne permet de prévoir que la forme sporadique *auricula* devait triompher de *auris* et aboutir aux formes telle qu'en français *oreille*, en italien *orecchia*. Pour les linguistes qui étudient l'histoire des langues romanes, c'est une grande commodité que d'avoir l'ancien latin ; les linguistes

qui étudient l'histoire des langues germaniques ou celle des langues indo-européennes en général, ont bien des raisons de la leur envier. Mais pour les romanistes comme pour tous les autres, le seul instrument de preuve valable est la comparaison des langues constituant le groupe étudié.

Dès lors, si, au lieu de comparer quinze ou vingt dialectes fortement distincts les uns des autres, mal localisés et dont les rapports mutuels sont mal connus ou inconnus, on peut comparer des centaines de parlers très voisins, ayant conservé d'une manière certaine leur ancienne répartition géographique, on dispose d'un moyen de recherches bien supérieur, et l'on a chance de déterminer avec une précision toute nouvelle l'histoire de l'ensemble des parlers étudiés.

Le livre de M. Gilliéron sur le nom de l'*Abeille* montre combien sont complexes les développements linguistiques. Le linguiste qui se contente de constater que *abeille* remonte à un latin *apicula* et que la forme est empruntée à des parlers de la France méridionale n'a aucune idée de la complication des faits qui ont déterminé la répartition actuelle des noms de l'abeille en France. C'est de constatations de ce genre qu'on est obligé de se contenter sur la plupart des domaines linguistiques. Le mérite de M. Gilliéron est de laisser entrevoir que, là même où l'on ignore trop de faits pour apercevoir le développement réel, il ne faut pas s'imaginer que, avec des correspondances simples, on touche la réalité des choses.

Précieux par les conclusions définitives qu'il apporte, le livre de M. Gilliéron l'est bien plus encore par la méthode qu'il enseigne. Jusqu'ici, on a fait la grammaire comparée des langues romanes avec la méthode qu'imposent les conditions de fait, assez fâcheuses, où se trouve la grammaire comparée des langues indo-européennes : aucune donnée sur la langue initiale; un petit nombre de groupes de langues très différentes entre elles; ces quelques données permettent surtout de saisir certaines grandes lignes, le détail n'étant perceptible que dans quelques rares cas spécialement favorables. Cette situation a eu, il est vrai, quelques avantages : elle a obligé à pratiquer la méthode comparative dans

sa pureté et à embrasser d'un coup de larges développements poursuivis durant de longs siècles. Mais elle a eu ses inconvénients : on a observé des correspondances plus ou moins régulières entre des langues ; mais ces correspondances résultent de successions complexes de changements dont le détail reste inconnu ; ce n'est qu'en observant des cas où le grand nombre des développements distincts permet de restituer le détail de ces successions qu'on arrive à comprendre comment se sont réalisées les correspondances observées.

Si l'on avait fait il y a cent ans le relevé des parlers locaux qu'on fait aujourd'hui, on aurait trouvé bien des faits curieux que l'extension du français commun a effacés déjà. Il faut se hâter de relever tout ce qui subsiste encore ; la disparition est plus rapide d'année en année.

Du reste, on est tout surpris de voir à quel point l'histoire de la langue française a été peu étudiée jusqu'ici et quelles lacunes apparaissent à qui veut, comme l'a fait M. Brunot avec tant d'autorité, en retenir l'ensemble. Faute de posséder des études assez poussées sur les divers groupes romans, la grammaire comparée générale des langues romanes n'est faite que sommairement. On est loin d'avoir tiré tout le parti possible des anciens textes latins qui reflètent à quelques égards l'usage vulgaire. Les textes d'époque mérovingienne n'ont pas été complètement exploités. Le seul dictionnaire de l'ancienne langue française qu'on possède, celui de Godefroy, est notoirement insuffisant [1]. Depuis du Cange, le vocabulaire des textes latins médiévaux qui offre si souvent des faits qui éclairent les langues vulgaires, n'a plus été systématiquement dépouillé. Faute d'utiliser les données qu'on possède, on a trop souvent étudié et exposé l'histoire du français, pour laquelle on a tant de faits, comme on fait de celle du latin ancien, pour laquelle on n'en a presque aucun. Les résultats obtenus par M. Gilliéron, grâce à une enquête systématique, montrent ce que l'on réalisera lorsque tout l'ensemble du travail sera organisé. Mais l'œuvre dépasse infiniment la force et les ressources de chacun des travailleurs isolés.

1. Je n'ai pu voir encore les fascicules parus du dictionnaire de M. Tobler.

SUR LE SENS LINGUISTIQUE

DE

L'UNITÉ LATINE [1]

Il y a une unité latine, mais elle n'est pas évidente au premier coup d'œil. Elle a besoin d'être dégagée, et il faut un petit effort d'attention pour en saisir la nature et la portée.

L'unité latine repose sur l'unité de l'empire romain.

Mais ses limites n'atteignent pas, à beaucoup près, celles de l'empire. Tout ce qui dans l'empire romain était de civilisation grecque est demeuré fidèle à l'hellénisme, sauf les cités grecques des côtes méridionales de l'Italie et des côtes de Sicile ; la séparation de l'Empire en Empire d'Orient, de langue grecque, et Empire d'Occident, de langue latine, répond à une réalité. L'unité latine recouvre seulement l'Empire d'Occident ; c'est ce qu'on peut appeler la *Romania*. La conquête arabe en a supprimé toute une partie, les provinces de l'Afrique du Nord, qui avaient reçu la langue latine. L'avance slave et albanaise dans les Balkans en a fait disparaître d'autres fragments ; mais l'influence romaine se manifeste encore par les nombreux emprunts que l'albanais a faits au vocabulaire latin. Ont subsisté les territoires actuels de l'Italie, de l'Espagne, du Portugal, de la France, une partie de la Suisse et de la Belgique, qui forment proprement l'unité latine. Les Roumains, ceux du royaume de Roumanie et ceux qui peuplent la Bessarabie autrefois annexée à l'empire russe et la Transylvanie autrefois annexée au royaume hongrois, sont aussi de langue *néo-latine* ou *romane* (les deux expressions sont équivalentes) ; mais ils ont pendant longtemps perdu le

1. *Revue des Nations latines*, I (1916), p. 179 et suiv.

contact avec les autres populations de langue latine, et par suite ils sont à part.

La partie occidentale de l'Empire avait une grande unité de civilisation qui se reconnaît à l'identité des institutions, des usages, des constructions officielles, et qui au moment de l'extension du christianisme a déterminé l'unité de l'Église. Ce qui reste d'unité latine vient de là.

Il ne suffit pas qu'il y ait un empire un pour déterminer une unité de langue, si puissant qu'il ait été, si grande qu'ait été son influence sur l'histoire du monde. L'empire achéménide n'a étendu sa langue à aucune des satrapies qu'il comportait, il n'a même pas étendu le parler perse aux tribus non perses du monde iranien. Une langue ne s'étend que si elle porte une civilisation douée de prestige. L'arabe n'a pu se répandre largement que parce qu'il portait avec lui une civilisation brillante, héritière des civilisations de l'Asie antérieure.

Le latin s'est répandu dans les parties de l'empire où on ne parlait pas le grec, c'est-à-dire dans toute la partie occidentale et au Nord des Balkans, parce qu'il portait avec lui la grande civilisation méditerranéenne que les Grecs ont propagée, à laquelle ils ont donné une originalité et une élévation singulières. Il est devenu la langue commune de la moitié occidentale de l'empire romain, parce qu'il apportait avec lui la culture romaine, et que la culture romaine, nourrie de la culture grecque dont elle avait pris tout ce qui était aisément accessible, mais dégagée de ce que cette culture avait de trop spécifiquement hellénique, représentait alors le plus haut degré de culture connu par l'humanité. Le latin apportait à l'Espagne, à la Gaule, aux provinces voisines du Danube, une civilisation faite pour l'humanité entière.

On sait bien que le latin a subi l'influence du grec. On ne se doute pas, en général, du point auquel, pour tout ce qui est le vocabulaire de la pensée et de la civilisation, les mots latins ont purement et simplement absorbé la valeur des mots grecs, comment, par exemple, avec une forme purement latine, un mot comme *cause* n'est au point de vue sémantique rien autre que le grec *aitia*, et ainsi toujours.

Le latin a de plus pris, même ailleurs, des éléments de civilisation plus ou moins importants. Le nom d'un vêtement long *tunica*, est pris au même mot sémitique auquel le grec a emprunté *khitôn*. Les Gaulois semblent avoir eu un art développé de la construction des chars ; César a encore trouvé le char de guerre en usage chez les Bretons de Grande-Bretagne, et il en a été gêné au premier abord ; aussi les Romains ont-ils pris aux gaulois des mots relatifs aux chars : *carrus*, *carpentum* sont des mots gaulois, et ce n'est pas au fait qu'ils habitent le sol de la Gaule, mais au fait que les Romains ont adopté ces mots que le français doit d'avoir *char* et *charpentier*. Langue de civilisation, le latin a pris de divers côtés les mots utiles à la culture.

Bien entendu, la question de race n'intervient pas. Rome a donné peu à peu le droit de cité romaine à tous les habitants de l'Italie : elle a fait de tous les habitants libres de l'Empire des citoyens romains. Nulle part elle n'a exterminé les habitants ; elle s'est contentée de leur apporter la civilisation romaine qui, dans la plupart des régions, après la guerre de conquête, a été acceptée volontiers et qui, une fois introduite, s'est développée normalement. Les habitants actuels des diverses parties de la *Romania* sont donc pour la plus grande partie les descendants des anciens occupants de ces territoires. La diversité des types ethniques est grande dans les pays de langue néo-latine, d'un pays à l'autre, d'une partie d'un pays à une autre partie, souvent d'un canton à l'autre. Il n'y a donc, entre les individus parlant les langues romanes, aucune communauté de « race ». Tout au plus faut-il noter qu'il s'est produit entre eux des mélanges.

Il ne subsiste de l'unité latine qu'une marque nette : les diverses langues parlées par les populations dites latines sont toutes des transformations du latin qui se parlait dans la *Romania*. Sans méconnaître l'importance de ce qui subsiste des anciennes institutions romaines chez les peuples latins, on peut mesurer leur unité à ce qui demeure d'unité dans l'ensemble des langues romanes.

L'unité linguistique de la *Romania* est brisée depuis longtemps en ce sens que le latin a pris en chaque région des formes diffé-

rentes et que les sujets parlants ont cessé de pouvoir se comprendre entre eux. Le français, l'italien, l'espagnol, le portugais et le roumain sont devenus autant de langues distinctes, et un Italien par exemple ne comprend pas immédiatement un Français, ou inversement.

Mais, malgré cette diversité certaine, ces langues ont conservé ou développé beaucoup de traits communs grâce auxquels l'ancienne unité transparaît et se continue : ce sont ces traits qu'il convient de mettre en évidence.

On laissera de côté le roumain, qui s'est trouvé de bonne heure placé sous l'influence de civilisations autres que le reste des langues de la *Romania* et qui par suite a divergé beaucoup d'avec les autres langues romanes. Si évident que soit le caractère latin du roumain, si forte que soit l'empreinte romaine en Roumanie, les conditions ont été là trop différentes pour que les observations présentées dans cet article s'appliquent toutes exactement au roumain. La situation à part de la Roumanie se reconnaît dès l'abord à ceci que les Roumains se sont rattachés non à l'Église d'Occident, mais à l'Église d'Orient. L'influence française, qui a été grande au XIX^e^ siècle, a heureusement relatinisé le roumain depuis quelques décades.

*
* *

La ressemblance des langues romanes entre elles est encore très sensible, et, même sans être linguiste, on aperçoit immédiatement que l'italien, l'espagnol, le portugais et le français se ressemblent plus entre eux qu'ils ne ressemblent à l'allemand, au suédois ou au danois d'une part, au russe, au polonais ou au serbe, de l'autre.

Cette ressemblance, très appréciable encore aujourd'hui, tient d'abord à ce que la date à laquelle les langues néo-latines ont commencé à diverger n'est pas très ancienne. Au III^e^ siècle après J.-C., quand la civilisation antique helléno-latine s'est mise à décliner et a été près de disparaître pour la première fois, la manière courante et familière de parler le latin sur laquelle reposent les langues romanes était, à quelques détails près, la

même sur toute l'étendue de l'Empire. A partir du moment où, vers le Ve siècle, l'Empire s'est définitivement brisé, le développement est devenu autonome dans chaque province, presque dans chaque localité.

En un espace de quinze cents ans toute langue change, et il est naturel que les langues romanes actuellement parlées diffèrent beaucoup du latin. Toutefois, la rapidité du changement n'est pas égale partout. Nulle part les changements n'ont été plus rapides ni plus radicaux que sur le domaine roman. Le français d'aujourd'hui diffère beaucoup plus du français du XIe siècle que le russe actuel ne diffère du russe de la même époque. Et, bien que la séparation ne soit pas beaucoup plus ancienne, les langues romanes sont bien moins pareilles entre elles que ne le sont les langues slaves par exemple. C'est que, presque partout où le latin se parlait au moment où l'empire romain s'est fragmenté, il y était d'importation récente, qu'il y succédait à d'autres langues ou à d'autres dialectes et que la tradition n'avait pas la fermeté d'une tradition linguistique établie depuis de longs siècles dans une population stable ; si en Gaule par exemple, les classes dirigeantes, les habitants des villes ne parlaient sans doute que latin au IVe siècle, il semble bien que le gaulois n'avait pas disparu partout à la campagne. C'est aussi que, au moment des invasions barbares et à la suite de ces invasions, il y a eu de grands mouvements de populations et de grands changements sociaux ; la civilisation urbaine, par exemple, a fait place presque partout à une vie essentiellement rurale. Des changements de cette sorte ont pour conséquence nécessaire une rapidité très grande des transformations linguistiques.

Il en est résulté que, entre le IIIe et le IXe siècle, par exemple, le latin a entièrement changé de caractère sur toute l'étendue de la *Romania*, et que, comme ces changements avaient lieu d'une manière indépendante dans chacune des parties du domaine, ils y différaient dans le détail. Aussi la forme particulière qu'a prise le latin varie-t-elle beaucoup de province à province. Alors que, sur toute l'étendue de l'empire russe, la langue est encore une presque comme au XIe siècle et qu'un paysan russe tout à fait

illettré comprend partout un autre paysan russe, il suffit parfois, sur le domaine roman, de quelques dizaines de kilomètres pour que des gens parlant leur patois local cessent de s'entendre.

Néanmoins le fait subsiste que les langues romanes divergent depuis une date relativement peu ancienne, et ceci suffit pour qu'elles aient gardé un air de parenté.

*
* *

La langue qui s'est brisée en langues diverses suivant les provinces n'était pas le latin de civilisation. Quand la langue s'est différenciée, c'est que la civilisation antique était brisée. La langue qui a subsisté dans l'usage parlé n'est plus le latin savant, littéraire, celui des gens cultivés ; c'est celui du peuple. Sans doute, cette langue du peuple est pleine d'influences savantes : partout, la langue populaire se nourrit de formes savantes ; elle est moins arrêtée qu'on ne croit souvent ; elle prend aux formes cultivées une grande partie des ressources qu'elle emploie. Mais elle met sa marque sur ce qu'elle emprunte. Si plein qu'il soit de mots et de formes dont l'origine est savante, le latin commun dont les transformations ont fourni les langues romanes, est une langue du peuple.

Le caractère vulgaire du vocabulaire roman se reconnaît à bien des traits. Ainsi le nom de la « maison » n'est plus le mot courant *domus*, c'est le mot *casa* qui désigne la petite maison populaire.

Une histoire curieuse est celle de l'adjectif *bellus* ; au point de vue latin, c'est un diminutif du même mot *duenos*, qui est devenu *bonus* ; et ce diminutif appartenait, comme la plupart des mots de ce genre, à la langue familière ; il désignait ce qui est gentil, agréable, et la nature du sentiment qui s'y attachait n'avait souvent rien de favorable ; un homme trop soigné, trop parfumé était qualifié de *bellus*. Or, c'est le mot qui a survécu en français pour indiquer ce qui est le plus élevé au point de vue esthétique et même moral. Par le fait que le parler populaire est devenu la langue de tous et qu'il a ainsi fourni la langue des hommes les plus cultivés, les termes qu'il comprenait ont pris une noblesse, une élévation qu'ils n'avaient pas.

Les noms des diverses parties du corps marquent bien le caractère vulgaire du vocabulaire roman. Le mot *bucca*, qui est resté pour désigner la « bouche », était nettement un mot vulgaire ; le mot latin commun était *ōs*. Pour désigner l' « oreille » les langues romanes n'ont pas gardé le mot latin *auris*, mais son diminutif *auricula*, familier comme la plupart des diminutifs et dont le caractère vulgaire se reconnaît à ce que c'est le mot employé dans un des rares textes vraiment vulgaires, celui qu'on lit dans les *tabellae deuotionis* écrites par des sorciers de bas étage.

Le nom français de la « tête » est *testa*, dont le sens ancien est « tesson » ; on voit tout ce qu'il y a de vulgaire dans le procédé ; c'est ainsi que le français populaire parle de la *fiole* ou de la *bobine* des gens, pour désigner leur « visage ».

Pour « parler », on se sert soit de *parabolare* (français *parler*, etc.), soit de *fabulare* (esp. *hablar*, etc.) ; c'est ce qui arrive quand, en français populaire, on dit : « Qu'est-ce que te racontes ? », pour « Que dis-tu ? ». Si l'on remonte en arrière, *parabola*, dont le verbe *parabolare* est dérivé, est un mot très savant ; c'est un terme de la rhétorique grecque, passé tel quel en latin ; les demi-lettrés qui ont écrit l'Évangile s'en sont servis pour nommer les enseignements donnés par Jésus sous forme de récits, de comparaisons ; et de la langue du christianisme le mot a passé dans l'usage populaire.

Les formes grammaticales des langues romanes procèdent en partie de tours populaires expressifs : c'est de la tendance populaire à s'exprimer avec force que proviennent des tours comme *j'ai à dire* (en français *je dirai*) au lieu de l'ancienne forme du futur ; on a *ecce iste* (vieux français *cest*, français *cet*) au lieu de *iste* « celui-ci ». Les exemples sont nombreux.

*
* *

L'indépendance des développements n'exclut pas le parallélisme. A beaucoup d'égards les déviations de l'usage latin qu'offrent les langues romanes, tout en étant identiques entre elles ou de même type, reposent non sur des innovations déjà réalisées en latin dans l'usage courant, mais sur des tendances

communes, sur des développements parallèles, qui ont eu lieu indépendamment dans chacun des parlers romans.

Le français et le provençal offrent de ce grand fait un exemple saisissant. L'innovation grammaticale la plus grave par laquelle les langues romanes se distinguent du latin ancien est la ruine du système de la déclinaison du nom, qui a été accompagnée universellement d'un renouvellement complet dans la manière de grouper les mots, de l'emploi de nombreuses particules qui se sont substituées aux formes casuelles défaillantes et de la fixation de l'ordre des mots, grâce à quoi l'on a rendu en partie ce que le latin rendait par ses formes casuelles. Or, cette ruine du système de la déclinaison n'est achevée ni en vieux français ni en vieux provençal. Dans ces deux dialectes, il ne subsiste plus six cas distincts comme en latin classique. Mais les masculins distinguent, au singulier et au pluriel, un cas sujet et un cas régime : au singulier le nominatif *caballus* a donné le cas sujet, vieux français *chevaus*, vieux provençal *cavals*, et l'accusatif *caballum* a donné le cas régime, vieux français *cheval*, vieux provençal *caval* ; au pluriel le nominatif *caballi* a donné le cas sujet, vieux français *cheval*, vieux provençal *caval*, et l'accusatif *caballos* a donné le cas régime, vieux français *chevaus*, vieux provençal *cavals*. Au contraire, toutes les autres langues romanes n'ont, depuis l'époque la plus ancienne, qu'une seule forme à chaque nombre, et, par suite, toute trace de déclinaison y a disparu. Le développement propre du français et du provençal a abouti finalement à la même simplification, et dès le XIV^e^ siècle, le français et le provençal ont perdu les derniers restes de déclinaison, ne gardant au singulier comme au pluriel que le cas régime, ce qui a abouti à l'opposition d'un singulier, français *cheval*, provençal *caval*, et d'un pluriel, français *chevaus* (écrit *chevaux*), provençal *cavals*. Des types exactement pareils se sont donc réalisés finalement en francais et en provençal, d'une part, dans les autres langues romanes, de l'autre ; mais ç'a été de manière tout à fait indépendante, et en conséquence surtout de l'identité des conditions initiales du développement.

Du reste la parité de l'italien, de l'espagnol et du portugais,

et aussi du roumain, dans l'abandon total de la déclinaison a été obtenue d'une manière indépendante dans chacune de ces langues. La forme du pluriel, — cas sujet et cas régime à la fois — reposant en italien et en roumain sur l'ancien nominatif latin *caballi*, est en italien *cavalli* et en roumain *cai* ; au contraire la forme unique de l'espagnol et du portugais repose sur l'ancien accusatif : espagnol *caballos*, portugais *cavallos*. Ici encore, si les résultats sont semblables — quoique non identiques dans le détail —, la ressemblance ne provient pas de l'identité du point de départ mais du parallélisme des développements.

Il ne serait sans doute pas aisé de trouver un exemple plus frappant de ces innovations parallèles qui ont donné aux langues romanes des structures si pareilles les unes aux autres. Mais les faits de ce genre sont innombrables. Se développant dans des conditions semblables, les langues romanes ont, en une large mesure, évolué dans une même direction.

*
* *

Dans les innovations des langues de peuples civilisés, les changements spontanés dont il vient d'être question ne sont pas les seuls, ni toujours les plus importants. La langue change aussi, pour une large part, sous des influences savantes. Le vocabulaire et la syntaxe sont soumis à ces influences.

Or, à l'exception du roumain, la grande influence savante qu'ont subie les langues romanes depuis le début, et presque sans interruption, a été celle du latin écrit, langue de la littérature, qui est restée la langue de l'Église, et, par là, la langue de toute la civilisation du moyen âge : on sait que, depuis l'antiquité jusqu'au dix-septième siècle, le latin a été pour toutes les nations de l'Europe occidentale la langue savante par excellence. Durant cette longue période les savants de l'Europe occidentale ont eu ainsi un moyen de communication qui était très éloigné de l'usage courant, qui avait l'inconvénient de les tenir loin des réalités et de les confiner dans des abstractions, de les priver ainsi de ce sentiment délicat des nuances qui n'existe que dans une langue vivante et familière, mais qui leur permettait de s'entendre

directement entre eux sans traductions, sans intermédiaires et sans apprentissage des langues de leur temps.

Le latin écrit qui, au prestige d'être la langue de toute la science, joignait celui d'être la langue de la religion, ne pouvait manquer d'avoir sur la langue parlée une grande influence. Cette influence s'exerçait d'autant plus aisément sur le domaine roman que, la langue parlée étant une transformation du latin, on saisissait souvent la parenté entre les mots vulgaires et les mots savants et que, si l'on venait à faire un emprunt, le mot pris au latin savant s'harmonisait presque toujours avec ceux que la langue courante avait conservés en les modifiant à sa manière.

Les emprunts des langues romanes au latin n'ont jamais cessé. Dans une langue qui, comme l'italien, a conservé aux mots un aspect phonétique assez proche de l'aspect latin, il n'est même pas toujours possible de discerner si un mot donné a subsisté continûment dans le langage courant depuis l'antiquité ou s'il a été pris à la langue écrite ; en français au contraire, le départ est généralement facile. Le mot *siècle* par exemple, bien que très ancien en français, se dénonce par sa forme comme ayant été pris par le français à la langue latine écrite dès avant les plus anciens textes écrits en français, et il a subi une forte adaptation au type vulgaire. On en peut dire autant de l'espagnol *siglo*, du portugais *seculo*. L'italien *secolo*, beaucoup plus près de l'original latin *saeculum*, ne se dénonce pas aussi évidemment comme un terme pris à la langue écrite et dû à l'influence de l'Église, bien qu'il ait aussi certainement pareille origine.

Par suite de l'action qu'a exercée la langue savante universelle qui était le latin, tout le vocabulaire abstrait du français, de l'espagnol, du portugais et de l'italien est composé pour ainsi dire de termes empruntés à cette langue ; et, bien qu'ils soient adaptés à chacune des langues où ils figurent, l'identité des termes dans les quatre langues se reconnaît au premier coup d'œil. Ainsi un même mot latin de la langue savante, comme *combinatio*, a été pris par les quatre langues, et l'on a en français *combinaison*, en italien *combinazione*, en portugais *combinaçao* et en espagnol *combinaçion* ; malgré de menues différences de prononciation, l'identité des quatre mots ne peut échapper à personne.

Le mot du latin parlé *causa* s'est maintenu dans les quatre langues ; il y a pris des formes qui en italien et en espagnol, sont à peu près identiques ; italien et espagnol *cosa*, portugais *cousa*, tandis que le français *chose* diverge fortement ; mais le mot savant *causa*, pris en un sens philosophique ou juridique, se reconnaît partout ; italien, portugais et espagnol *causa*, français *cause*. Or, ce mot n'est pas resté confiné dans des langues savantes plus ou moins techniques ; il est devenu un mot de la langue courante, et rien n'est plus familier que *cause* en français d'aujourd'hui.

Une très grande partie, la plus frappante peut-être, de l'unité qu'offrent actuellement les langues romanes, tient à ce qu'elles ont pris au même latin écrit, à la langue de la philosophie, de la théologie, du droit au moyen âge, leur vocabulaire abstrait et à ce que ce vocabulaire a été assimilé au vieux fonds indigène, aussi d'origine latine, à peu près dans les mêmes conditions. Comme la presse et l'école grandissent sans cesse l'influence de la langue écrite sur la langue parlée, le vocabulaire abstrait, commun à toutes les langues romanes occidentales, se répand de plus en plus.

Dans la mesure où le vocabulaire savant n'est pas pris au latin, il l'est au grec dans des conditions qui sont les mêmes sur tous les domaines : on a en italien, en portugais et en espagnol *telescopio*, *telegrafo*, et en français *télescope*, *télégraphe*, etc. Tout comme les mots abstraits pris au latin écrit, ces termes entrent dans le langage courant et y prennent une large place : *auto-* dans le mot hybride, mi-grec, mi-latin, *automobile* vient de ce que cet élément *auto-* est entré dans la langue courante ; et le mot *automobile*, avec sa forme abrégée *auto*, est devenu l'un des plus usuels de la langue, tout comme le mot, également hybride, *bicyclette*. Il se trouve ainsi que, par là encore, les langues romanes occidentales tendent actuellement à se rapprocher, au lieu d'accroître les différences qui les séparent.

C'est à l'unité de civilisation de l'empire romain que l'unité romane devait son existence. C'est à l'unité de culture de toute l'Europe occidentale au moyen âge et jusqu'à l'époque moderne

que les langues romanes doivent d'avoir sans cesse consolidé cette unité, malgré la divergence des développements linguistiques. Tant il est vrai que toute unité linguistique repose sur une unité de civilisation.

*
* *

Les histoires de France, d'Italie, d'Espagne, de Portugal n'ont pas cessé de se mêler ; tantôt alliées et tantôt ennemies, les populations des quatre pays ont eu constamment des relations. Soit par la vie sociale, soit par les arts et la civilisation, chacun d'eux a pris tour à tour une avance à quelque point de vue. Tous ces rapports ont eu pour conséquence des emprunts de mots d'une langue à l'autre. Le français, par exemple, comprend une foule de mots italiens et espagnols.

Ainsi la banque est une institution d'origine italienne, et les Lombards en ont fourni le modèle à toute l'Europe. Le mot *banque* en français, le mot *banco* en espagnol sont des emprunts à l'italien *banco*. Le mot *crédit* en français, le mot *credito* en espagnol sont, de même, pris à l'italien *credito*. Les conquêtes coloniales des Espagnols au seizième siècle leur ont fait connaître le chocolat, dont ils ont appris le nom des Mexicains. Le *chocolate* espagnol a donné le mot *chocolat* au français, et *cioccolata* à l'italien.

Ce sont là quelques exemples pris au hasard pour montrer à quel point des termes de civilisation d'usage courant ont passé d'une langue romane à l'autre et combien les mots ainsi empruntés sont demeurés semblables entre eux.

Les emprunts de mots, pour nombreux et significatifs qu'ils soient, ne donnent pas la mesure de l'importance des relations linguistiques entre les langues romanes. La littérature française a agi sur les débuts de la littérature italienne au moyen âge ; la littérature italienne a fourni leurs modèles aux écrivains français du seizième siècle ; et, au début du dix-septième siècle, la littérature française subit encore une forte influence italienne, espagnole et même portugaise tout à la fois ; puis les Français prennent une avance et ce sont eux qui, au dix-septième et au

dix-huitième siècle, fournissent plutôt des modèles. Au cours de toutes ces actions successives et réciproques, les mots se sont imprégnés de sens étrangers, les tours de phrases ont reçu des équivalents. Les vocabulaires des langues romanes occidentales, tout en étant composés de mots dont l'aspect extérieur varie sensiblement, se trouvent ainsi exprimer assez exactement les mêmes notions, de sorte que, dans une large mesure, ils se traduisent avec précision les uns les autres. A parcourir un journal italien, espagnol ou portugais, le lecteur français a souvent l'impression de retrouver sa propre langue sous un vêtement un peu différent, tant les mots s'équivalent, tant les expressions se traduisent exactement, tant les tours de phrases sont parallèles.

*
* *

L'unité latine, qui se manifeste par des faits de langue, est réelle. Mais, on le voit, elle n'est pas chose simple, et elle résulte de circonstances très diverses. Telle qu'elle est, elle a une grande portée, parce qu'elle résulte de la vieille unité de leur civilisation, qui ne s'est jamais tout à fait brisée, et parce qu'elle permet aux gens de langues française, italienne, espagnole et portugaise de s'entendre entre eux avec très peu d'effort, qu'elle les met pour ainsi dire de plain pied, et d'autant plus de plain pied qu'ils sont plus des hommes civilisés.

Toute unité linguistique traduit une unité de civilisation soit présente, soit passée. L'abolition de l'unité traduit la rupture de l'unité de civilisation : s'il y a aujourd'hui des langues romanes diverses, c'est parce que la civilisation antique s'est ruinée du IIIe au VIIIe siècle après J.-C. Et toutes les réactions contre la différenciation des langues traduisent un effort fait pour maintenir l'unité de civilisation ou pour y revenir : l'unité romane actuelle tient avant tout à l'unité profonde de la civilisation européenne occidentale.

LA RELIGION INDO-EUROPÉENNE[1]

Il peut sembler étrange, au premier abord, qu'un linguiste, qui n'est rien que linguiste, se sente qualifié pour traiter de la religion indo-européenne.

La raison est simple : c'est que les peuples de langue indo-européenne n'ont connu l'écriture que très tard, à un moment où elle était pratiquée à Babylone et en Égypte depuis beaucoup de siècles ; seuls de toutes les langues indo-européennes, le sanskrit, l'iranien, le grec et les dialectes italiques, dont le latin est le principal, sont attestés avant l'époque chrétienne ; toutes les autres langues, slave, baltique, germanique, celtique, arménienne, ne sont attestées qu'après le IVe siècle après Jésus-Christ — en partie beaucoup après — et par des textes chrétiens. Or, plus que toute autre histoire, l'histoire des religions a besoin de textes, et de textes conçus dans la langue même du peuple étudié. Il est donc impossible de faire l'histoire ancienne des religions pour les peuples de langue indo-européenne.

La linguistique n'a pas le moyen de remplacer l'histoire, et elle n'y prétend pas ; mais, à défaut des précisions que fourniraient des textes suivis, elle apporte au moins les seules lueurs qu'on puisse espérer sur une période qui ne nous a laissé aucun témoignage direct ni de ses actes ni de sa pensée. Toutefois, avant d'interroger la linguistique, il importe de déterminer exactement ce qu'on en peut attendre ici.

I

Il est bien connu que certaines langues de l'Asie, et presque toutes les langues de l'Europe, appartiennent à un même groupe, qu'on est convenu d'appeler groupe indo-européen, les Allemands

1. Extrait de la *Revue des idées*, IV, p. 689 et suiv.

disent indo-germanique ; le nom est arbitraire, et n'a pas d'importance.

Dire que le sanskrit, le perse, le slave, le germanique, le celtique, le grec, l'arménien sont des langues du groupe indo-européen, c'est affirmer que ces langues sont des transformations diverses d'une seule et même langue. De cette langue, on n'a aucun témoignage direct, puisqu'elle n'a pas été écrite. Mais l'identité d'origine des langues en question se traduit par certaines ressemblances ; et ces ressemblances ne sont pas capricieuses et fortuites ; comme le développement des langues est soumis à des lois, il existe des systèmes réguliers de correspondances de chacune des langues attestées avec toutes les autres langues du même groupe. On appelle langue indo-européenne, ou, tout simplement, l'indo-européen, l'ensemble de ces systèmes de correspondances.

En fait, de la langue dont les correspondances qu'étudie la grammaire comparée supposent l'existence, on ne sait rigoureusement rien. On ne sait par quels hommes elle était parlée : des hommes qui parlent aujourd'hui les diverses langues indo-européennes, les uns sont dolichocéphales et les autres brachycéphales, les uns sont bruns et les autres blonds, les uns sont grands et les autres petits, et, à vrai dire, la plupart sont de races très mélangées ; par quelle race était parlée la langue indo-européenne, on l'ignore, et l'on ignore de même si c'était par une race pure ou non. — On ne sait où cette langue était parlée ; on l'a localisée en Asie, puis on l'a transportée sur toutes sortes de régions de l'Europe, des steppes de la Volga aux Carpathes, des Carpathes à la Scandinavie et de la Scandinavie à la Lituanie ; et les raisons décisives manquent pour s'arrêter nulle part. — On ne sait quand cette langue a été parlée ; on voit bien dans quelle mesure les divers dialectes divergent entre eux ; mais comme le degré de rapidité avec lequel les langues se transforment est très variable, on n'en saurait tirer aucune conclusion, même approximative.

L'indo-européen est donc un système de correspondances linguistiques, qui suppose une langue x parlée par des hommes x en un lieu x, en un temps x.

Ce système de correspondances a une grande importance en ce qu'il permet de déterminer quels rapports soutiennent entre elles les langues de la famille et d'entrevoir au moins les traits principaux du développement de ces langues dans les siècles qui ont précédé immédiatement l'époque des plus anciens textes. On prolonge ainsi l'histoire de chacune des langues indo-européennes, au moins dans une certaine mesure, jusqu'à la période où toutes ces langues n'en faisaient qu'une. Il est bien inutile d'ajouter aujourd'hui que cette période n'est pas une période *primitive*, et que la grammaire comparée ne nous rapproche pas sensiblement de l'origine du langage : l'indo-européen n'est pas plus « primitif » que le vieux babylonien ou le vieil égyptien, qui ne le sont absolument pas.

Les seules correspondances à examiner ici sont celles qui nous renseignent sur des faits de vocabulaire. Il est clair que, si une correspondance atteste dans le système linguistique l'existence d'un mot désignant une notion définie, on devra attribuer cette notion au peuple qui parlait la langue supposée par le système. Mais les conclusions de ce genre sont nécessairement très vagues : les notions attachées aux mots changent souvent sans que les mots changent pour cela. D'ailleurs, le nombre des faits utilisables est petit : car on n'a de correspondances que pour les termes les plus généraux ; tout ce qui est technique et particulier est spécial à chacune des langues du groupe et ne comporte de correspondances dans aucune autre. La grammaire comparée ne peut donc fournir sur l'état de civilisation du peuple supposé par la langue indo-européenne que des indications vagues, incomplètes et souvent douteuses ; on s'est longtemps fait à cet égard des illusions dont on est revenu aujourd'hui.

Si peu qu'on doive espérer, il vaut la peine d'interroger la linguistique indo-européenne et d'examiner ce qu'elle peut enseigner sur les idées religieuses du peuple qui a fourni les langues employées par une partie notable de l'Asie, et par presque toute l'Europe. En procédant avec critique, on peut obtenir des conclusions peu nombreuses et qui ne répondent guère aux vastes espérances qu'on a conçues autrefois : la grammaire comparée

est devenue une science puissamment développée, tandis que la mythologie comparée, fondée sur la linguistique, peut toujours tenir à l'aise en un exposé de quelques pages. Mais ces quelques pages méritent d'être écrites, parce qu'on y peut consigner des résultats précis, et dont rien autre ne saurait fournir l'équivalent.

II

Une correspondance nette, l'une des plus claires et des plus sûres qu'on possède dans tout le vocabulaire indo-européen, établit l'existence d'un terme pour l'idée de « divinité »; c'est sanskrit *devaḥ*, lituanien *dëvas*, vieux prussien *deiws* (génitif *deiwas*), latin *deus* (vocatif *dīue*, etc.), vieil irlandais *dia*, gaulois *dêvo-* (dans *Devognata*, « née d'un dieu », nom propre), vieil islandais *tívar* (pluriel).

Cette correspondance s'exprime par le symbole **deiwos*, qui donne au moins une idée de ce qu'a pu être le mot représenté par les formes citées. On voit que le terme n'est pas limité à une partie du domaine indo-européen ; on le rencontre aux deux extrémités orientale et occidentale du domaine ; s'il ne figure pas en slave, il existe dans les dialectes baltiques qui lui sont étroitement apparentés ; et, s'il n'est pas conservé en grec, il y a du moins un dérivé également ancien : grec ἔϊος, identique à sanskrit *divyaḥ* « céleste » et à latin *dius* « divin ».

Si donc on a quelque part le droit de parler d'un mot indo-européen, c'est ici. La forme est partout celle que font attendre les règles de correspondances phonétiques : ces règles écartent absolument l'idée que le grec θεός (*theos*) aurait rien à faire avec latin *deus* ; on a fait sur l'origine de θεός diverses hypothèses, dont l'une, celle d'après laquelle θεός aurait signifié à l'origine « souffle, esprit », n'est pas inadmissible et semble même vraisemblable ; toute affirmation sur ce point est impossible : on ne sait certainement qu'une chose, c'est que θεός ne répond pas au latin *deus* : en matière d'étymologie, comme en toute matière scientifique, les premières apparences sont trompeuses.

Le sens de **deiwos* est « divinité » de la manière la plus géné-

rale, et ce sens est le même partout ; après avoir désigné d'une manière générale les anciens dieux, le mot avait assez de souplesse pour s'appliquer au *Dieu* unique judéo-chrétien, et nous l'employons encore à la fois pour le *Dieu* chrétien et pour tous les dieux. Seul l'iranien diverge, mais pas d'une manière essentielle : les *daêva* de l'Avesta ne sont pas des dieux, mais ce sont des êtres non humains, les ennemis des dieux, les démons. Pour « dieu » il y a ici un terme dont l'extension ne va pas loin : *baga* du vieux perse ne se retrouve que dans le *bog* slave ; l'emploi ne diffère du reste pas de celui de **deiwos*. Il est inutile de parler du mot germanique (allemand *gott*), car il est isolé, et les hypothèses qu'on peut faire sur l'étymologie peuvent être plus ou moins plausibles, mais n'apprennent rien sur les faits indo-européens.

Si le mot **deiwos* était isolé, la linguistique n'enseignerait rien de plus que la simple existence de l'idée de « divinité », ce qui serait peu ; mais, par bonheur, le mot fait partie d'un groupe défini et connu, et l'on peut déterminer à quel genre de notions l'indo-européen a fait appel pour désigner la « divinité ».

Le sanskrit védique a un mot, *dyauḥ*, qui signifie tout à la fois « ciel » et « jour » ; ce mot a une déclinaison assez compliquée, conforme d'ailleurs aux règles générales de la flexion indo-européenne ; et à ce mot correspond un mot grec connu, fléchi de même qui est le nom d'un dieu particulier, mais visiblement d'un dieu du ciel, car c'est le dieu qui tonne, et qui pleut : dans grec ὕει « il pleut », *il* est Ζεύς (*Zeus*) : Ζεὺς ὕει « Zeus pleut » :

Nominatif : sanskrit *dyauḥ*, grec Ζεύς.

Datif : sanskrit *divé*, grec Δι-Fί.

Accusatif : sanskrit *dyām*, grec Ζῆν(α).

En latin, on retrouve le même vocable, mais sous des formes plus compliquées encore, car ce même mot a fourni plusieurs termes, dont on retiendra seulement les deux principaux. L'un tiré du nominatif, du vocatif et du locatif désigne un dieu particulier, qui est tout pareil à Ζεύς : *Juppiter*, c'est Ζεῦ πάτερ « ô *Zeus* père » ; *Joue*, c'est le locatif sanskrit *dyavi* ; l'osque a *Diuvei* à côté de *Iuvei*, et le vieux latin *Diouei Dioue* ; mais au sanskrit *dyām* répond lat. *diem* « jour », et cet accusatif a donné lieu à la

formation d'un mot lat. *dies* « jour ». Sans entrer dans plus de détails linguistiques, on voit qu'il existe un mot indo-européen symbolisé par **dyeus *dyem *diwes*, etc., qui signifie « ciel » et « jour ». Et ce mot doit évidemment être groupé avec **deiwos* « dieu » et **diwyos* « divin ».

Du coup la valeur étymologique de **deiwos* est précisée : le « dieu » indo-européen est un être lumineux, céleste ; le dieu, c'est le lumineux, le céleste. Et ce sens a dû rester longtemps dans la conscience même des sujets parlants ; en effet le finnois a emprunté aux dialectes baltiques le mot qui y est représenté par vieux prussien *deiws*, lituanien *dëvas* ; or, le mot finnois *taivas* ne signifie pas « dieu », mais « ciel ».

La notion du « dieu » indo-européen se laisse préciser encore et compléter par une autre remarque. L'idée de « dieu » n'est qu'une abstraction à laquelle on parvient en partie en extériorisant certaines notions d'expérience humaine, et en partie en niant d'autres notions d'expérience, de manière à créer une différence essentielle entre homme et dieu. L'homme étant un être terrestre et mortel, les dieux seront célestes et immortels : ces deux notions sont souvent indiquées chez Homère, par exemple ; on est ainsi amené à nommer l'homme par opposition, soit le « terrestre », soit le « mortel ».

Le dieu étant *céleste*, l'homme est *terrestre* ; de là l'un des principaux termes qui désignent l'homme dans les langues indo-européennes[1] : latin *homo*, osque *humuns* « les hommes », gotique *guma*, vieux haut allemand *gomo* (qui se conserve dans le nom allemand du « fiancé » *Bräuti-gam*), vieil islandais *gume*, lituanien (pluriel) *žmonės* ; ce mot est à rapprocher du nom de la terre : latin *humus*, grec χαμαί, lituanien *žẽmė*, vieux slave *zemlja*, zend *zəm-*, doublet de sanskrit *kṣam-*, grec χθών. Le mot *dieu* et le mot *homme* survivent encore en français ; le sens originel en est perdu ; seul le linguiste y discerne l'être céleste et l'être terrestre ; mais ces deux mots de notre langue actuelle sont les témoins de conceptions antiques, qui ont d'ailleurs en partie subsisté. Dieu n'est plus le ciel, mais il est toujours aux cieux.

1. Voir ci-dessus, p. 272 et suiv., sur le nom de l'homme.

Il est vrai que, dans notre conception moderne du monde, depuis Galilée et Copernic, cela ne signifie plus rien, et qu'un homme qui pense ne peut attribuer à « Notre père qui êtes aux cieux » aucun sens saisissable ; mais la terre a beau tourner, on a beau le savoir, on regarde toujours vers le ciel pour prier. Les vieux mots ont perdu leur signification et les vieilles idées leur valeur, mais les usages, les gestes subsistent.

Le dieu étant *immortel*, l'homme est *mortel* ; et Homère, qui qualifie les dieux d'ἄμβροτοι « immortels », nomme fréquemment les hommes βροτοί « mortels » ; ou avec une autre forme, plus claire, μορτοί ; de même l'homme se dit *mard* en arménien, et le sens de ce mot est perdu dès le début de l'époque historique : *mard* pour l'arménien est l'« homme », et non le « mortel » ; le vieux perse a de même *martiya*, et le persan moderne *mard*. Cette dénomination de l'homme par la racine *mer-*, qui signifie « mourir », ne se rencontre pas dans les dialectes occidentaux. Mais l'irlandais a une dénomination toute pareille : irlandais *duine* « homme » est un dérivé d'une forme apparentée à gotique *diwans* « mort », et à allemand *tod*. C'est toujours la même idée, si ce n'est pas le même matériel linguistique.

Enfin, pour Homère, les dieux sont les « donneurs de biens » et l'antiquité de cette notion est marquée par le nom iranien et slave ; le mot sanskrit *bhagaḥ* signifie à la fois « bien qu'on se partage », et « dieu qui partage » ; et c'est ce mot qu'on retrouve dans le vieux perse *baga* et dans le slave *bog* ; en slave même *ubog* et *ne-bog* signifient « pauvre » (celui qui n'a pas de part de richesse).

Telle est la conception du dieu indo-européen : céleste et lumineux, immortel, donneur de biens ; et cette conception n'est pas très éloignée de celle de l'homme du peuple dans l'Europe d'aujourd'hui.

III

Par sa nature et par les conditions mêmes où elle se présente, la linguistique ne peut enseigner plus que des notions très générales. Mais, pour compléter la théorie, il reste à relever quelques faits négatifs qui sont importants, et peut-être significatifs.

Tout d'abord l'absence d'un terme indiquant quoi que ce soit qui ressemble à un lieu de culte, à un instrument servant au culte, à un sacrifice en un mot. Il y a un mot indo-iranien qui signifie « sacrifier », sanskrit *yájati, yajate*, zend *yazaite* « il sacrifie » ; mais ce terme ne dépasse pas les limites de l'unique dialecte indo-iranien ; le seul terme qu'on ait trouvé à en rapprocher dans une autre langue, c'est le grec ἄζομαι, ἅγιος ; mais d'abord l'étymologie n'est pas évidente ; il est simplement possible, pour la forme, que sanskrit *yaja-* = grec ἁγε-, et ce n'est pas même probable, pour des raisons techniques, dans le détail desquelles on ne peut entrer ici ; et quant à la signification, qui importe avant tout, si le rapprochement est valable, il ne prouve rien, car le grec ἄζομαι ne se rapporte à aucun acte de sacrifice ; il indique plutôt un sentiment de crainte religieuse, de vénération, et en l'absence d'une troisième langue qui permette de décider entre le sens indo-iranien « sacrifier » et le sens grec « avoir une crainte religieuse », on ne peut rien dire du sens indo-européen. Il est probable que le latin *sacer, sanciō, sanctus* est en réalité le mot apparenté à ἅγιος du grec ; les sens concordent ; et la différence entre *c* de latin et γ du grec se laisse expliquer.

Ce n'est pas à dire que l'idée de « sacré », qui est l'idée fondamentale en matière religieuse, n'ait pas eu d'expression en indo-européen ; mais on n'a qu'une correspondance pour l'attester au point de vue linguistique, et cette correspondance ne s'étend qu'à un domaine linguistique très étroit : le baltique a le lituanien *szventas*, le slave a le vieux slave *svętŭ* (russe *svjat*), l'iranien a le zend *spəntô* ; on retrouve ici un mot limité à un petit domaine constitué par le baltique, le slave et l'iranien ; ceci rappelle le slave *bog* à côté de perse *baga*. On a aussi rapproché un nom germanique du « sacrifice », gotique *hunsl*, mais d'autres explications ont été proposées, et le sens diverge.

Ce n'est pas à dire non plus qu'il n'y avait pas d'hommes investis, au moins temporairement, de fonctions religieuses. Mais aucune correspondance linguistique bien établie n'en porte la trace. Le *brahman-* védique a souvent été rapproché du *flāmen* romain ; mais en admettant le rapprochement, qui n'est pas évi-

dent, en tout cas le mot ne se trouverait qu'en sanskrit et en latin : il s'agirait d'une de ces concordances de termes religieux entre l'indo-iranien et l'italo-celtique que M. Vendryes a mises en évidence et qui sont remarquables (voir l'article des *Mémoires de la Société de linguistique*, XX, p. 265 et suiv.). Il y a bien un rapprochement sûr, mais le mot qu'il fournit n'est celui d'aucune fonction religieuse définie ; de plus, il ne figure que dans les dialectes occidentaux : latin *uates* prophète, « devin », gaulois οὐάτεις « devins » traduit par grec μάντεις, irlandais *fāith* « poète », gotique *wods* « possédé, furieux », vieux haut allemand *wuot* « fureur » (l'allemand *Wuth*), vieil islandais *ōðr* « poésie » (de là le dérivé *Edda*) ; le personnage que désigne ce mot a certainement un caractère religieux, mais ce n'est pas proprement un prêtre, c'est un être inspiré, un devin.

Donc pas de mots sûrs pour désigner le lieu du culte, le sacrifice, ni le prêtre. Ceci s'explique aisément si l'on songe que de l'indo-européen on ne possède que les termes généraux, ceux qui étaient d'usage commun sur l'ensemble du domaine. Or, le culte de l'homme peu civilisé est essentiellement le culte d'une tribu : la tribu a ses lieux de culte, ses sacrifices et ses officiants qui lui sont strictement propres. Le manque de termes communs indique l'absence d'institutions communes ; et ce n'est pas un résultat de peu d'importance. Rien de plus particulier que le culte d'une population de civilisation inférieure.

Et dès lors on comprend pourquoi il n'y a pas de nom de dieu particulier qui soit indo-européen ; chercher à retrouver en indo-européen un nom de dieu, c'est supposer un culte commun à l'ensemble des populations dont la langue est désignée par le nom d'indo-européenne ; cette hypothèse n'est pas très vraisemblable a priori, et, le fût-elle, elle ne se vérifie pas. On s'est donné beaucoup de mal pour retrouver dans l'Inde les dieux helléniques et, en Grèce, les dieux hindous, et l'on n'a abouti à aucun résultat. Tous les rapprochements qu'on a proposés viennent se heurter aux lois strictes des correspondances phonétiques, et ceux des linguistes qui essaient de sauver les moins compromis de ces rapprochements en sont réduits à avouer que la phonétique

des noms de dieux n'est pas celle des autres mots, ou qu'il est intervenu des accidents impossibles à justifier dans le détail. *Indra* reste propre à l'Inde, *Apollon* à la Grèce, *Mars* à l'Italie, et ainsi de suite.

La linguistique ne fournit à la mythologie comparée presque aucun fait utilisable, et les illusions qu'on a pu avoir vers 1850 et que le talent de Max Müller a largement propagées ne sont plus partagées aujourd'hui par aucun des hommes compétents : il peut y avoir une mythologie comparée, mais elle ne sera pas fondée sur la linguistique, parce que la grammaire comparée ne fournit que des termes généraux, et que les cultes étaient particuliers.

IV

Il subsiste pourtant un grand fait, et où la linguistique est intéressée. L'archéologie préhistorique de l'Europe ne révèle guère d'idoles ; et partout où l'on a quelque témoignage sur les peuples de date indo-européenne ou en état de civilisation peu avancée, ces témoignages indiquent l'absence de dieux personnels. L'onomastique indo-européenne concorde avec ces constatations ; elle n'indique pas qu'une grande importance ait été attribuée à des dieux personnels : les noms propres de personnes indo-européens sont des composés à deux termes qui indiquent certaines qualités et non pas des dérivés de dieux personnels ; on peut s'appeler « qui a une bonne réputation », sanskrit *Suçravas-* zend *Husravah-*, gr. Εὐκλέης, Εὐκλῆς, mais on ne s'appelle pas en général « serviteur de tel ou tel dieu », comme dans les langues sémitiques. Il y a bien des noms comme Ἀπολλώνιος (Apollinien) en Grèce, ou *Esugenos* (descendant d'*Esus*) en Gaule, mais c'est une petite minorité ; le type courant est autre.

Et ceci conduit à une remarque importante : les seuls noms de personnages divins qui soient communs à plusieurs langues indo-européennes sont des noms d'astres, de phénomènes naturels, ou de choses de ce genre. Le plus bel exemple de cette série a déjà été cité : c'est le sanskrit *Dyauḥ pitá*, le grec Ζεύς, le latin *Juppiter* : on a ici un nom divin commun à plusieurs langues,

parce que l'adoration du ciel lumineux se retrouve sur tout le domaine indo-européen, ainsi que l'atteste à lui seul le mot **deiwos*. Seulement, il a pu arriver que ce nom ait été réservé au ciel en tant qu'être divin, et que le ciel matériel ait reçu un autre nom ; c'est ce que l'on observe en grec et en latin ; alors il n'est presque plus sensible que Ζεύς ou *Juppiter* soit le ciel, et l'on y sent de plus en plus un simple dieu personnel ; ce sont là des faits proprement romains ou grecs, et qui ne remontent pas à la période indo-européenne.

Là où l'ancienne religion est encore attestée par des témoignages un peu précis, on voit le soleil et la lune divinisés et adorés sous leur nom ; ceci apparaît aussi bien en Lituanie au XVI^e^ siècle — on sait que le paganisme s'est longtemps maintenu chez les Lituaniens — que dans l'Inde védique ou dans la Grèce ancienne. Et c'est sous leur nom courant que ces astres sont divinisés : tel est bien l'état indo-européen ; astres et phénomènes naturels sont divinisés, sous leur nom ordinaire [1].

Le grand dieu slave *Perunu* est le tonnerre, la foudre, et le sens de « tonnerre et foudre » est encore bien conservé en slave même ; la « foudre » est encore *piorun* dans le polonais d'aujourd'hui. Le lituanien a une forme un peu différente, *Perkúnas*, avec même valeur. Et ceci éclaire le nom du dieu védique de l'orage : *Parjanyaḥ ; Parjanyaḥ* n'a plus en védique le sens d'« orage » ; ce n'est plus qu'un personnage divin ; mais le slave et le lituanien révèlent le sens du nom, qui a été d'abord un nom commun, et n'est devenu un nom propre que par un développement secondaire. Ce nom se rattache à une racine **per-* signifiant « frapper » ; l'alternance d'un type **per-* avec un type **perk-*, **perg-* se retrouve, de manière remarquable, dans le verbe arménien dont l'aoriste est *hari* « j'ai frappé », et le présent, *harkanem* « je frappe » (le *k* arménien représente un ancien *g*) ; la gutturale se retrouve en celtique, dans irlandais *orgim* « je tue », par exemple. Il est probable que le grec Τρίτων était anciennement la « mer », sim-

1. On a vu ci-dessus, p. 215 et suiv., que des noms d'éléments tels que l'eau et le feu peuvent être de genre neutre, ou de genre « animé », suivant qu'on les considère comme des choses ou comme des agents, et que, s'ils sont considérés comme des agents, ils sont souvent divinisés.

plement ; or il rappelle de près *triath* « mer » (génitif *trethan*) de l'irlandais.

Et ce ne sont pas seulement les phénomènes naturels qui sont divinisés ainsi ; un bel exemple indo-iranien montre les faits sociaux divinisés de la même manière. L'un des personnages divins les plus importants de l'Inde, et de l'Iran est *Mitra* ; or, encore dans l'Avesta, *miθrô* désigne à la fois le « contrat » et le dieu *Miθra* ; dans l'Inde les choses sont un peu moins claires, mais encore transparentes : *mitraḥ* est l'ami, *mitram* l'amitié ; le rapprochement avec l'iranien montre assez qu'il ne s'agit pas, à l'origine, de l'amitié sentiment, mais du contrat d'amitié, tel que le pratiquent les hommes à un certain état de civilisation. *Mitra*, c'est le contrat ; le mot est apparenté au slave *mirŭ* qui signifie « paix » et à la grande famille de sanskrit *mayate* « il échange », lituanien *maĩnas* « échange », latin *com-mūnis* « commun », gotique *gamains* (allemand *ge-mein*) « commun ». En tant qu'il est le contrat divinisé, *Mitra* est celui qui apporte la sanction du contrat ; il voit donc tout, et même ce qui est caché ; il est l'œil qui voit tout ; et, comme le soleil est lui aussi l'œil du ciel, l'œil qui voit tout (en irlandais, le nom celtique du soleil, *súil*, est devenu ainsi le nom commun de l'œil), Mitra est rapproché du soleil, ce qui l'a fait prendre pour un mythe solaire, en dépit de l'évidence de l'étymologie qui révèle un fait social divinisé. Le sens de l'origine du dieu persiste encore dans une formule traditionnelle hindoue : l'hôte à qui l'on présente le plat d'hospitalité dit à ce plat : « je te regarde avec l'œil de Mitra ». Tout le sens de Mitra est là : et le contrat d'amitié, et le regard qui en résulte.

Il est inutile de multiplier les exemples : ceux qui ont été donnés suffisent pour révéler le caractère du dieu indo-européen : c'est un fait naturel ou social auquel on attache une importance particulière ; le dieu n'a pas un nom distinct de celui du fait. Le dieu n'est pas une personne ayant un nom propre ; c'est le fait lui-même, c'est son essence, sa force intime. Pour autant que la linguistique laisse entrevoir les choses, le culte indo-européen ne s'adresse donc pas à des êtres autonomes, mais aux forces naturelles et sociales elles-mêmes.

A PROPOS DE « QUALITAS[1] »

On se représente mal comment nous pourrions ne pas opposer la *qualité* à la *quantité*. Et il n'y a boutique si humble que la marchandise n'y soit de première *qualité*. Or, ce mot est tout étranger ; *qualité* n'est rien que le latin *qualitas*, et *qualitas* rien que le masque latin du grec ποιότης. Le mot est du petit nombre de ceux dont l'histoire est connue tout entière. Voici en bref cette histoire.

Pour Platon, le mot est nouveau. Le dérivé ποιότης, de ποιος, est employé près de l'adjectif dont il dérive, qui signifie « quel, lequel », pour indiquer le fait d'être « tel ou tel », tout comme ὁμοιότης « fait d'être semblable » est tiré de ὅμοιος « semblable », ἑτεροιότης ou ἀλλοιότης « fait d'être autre », de ἕτεροῖος ou ἀλλοῖος « qui a le caractère d'être autre, d'être différent ». Platon recourt à ces mots nouvellement fabriqués dans ses dialogues techniques, dans le *Timée*, dans le *Parménide*, dans le *Théétète*, là où il joue avec les idées, où il n'est le poète que des idées. Qu'on en juge par la phrase du Théétète (182 *a*), où figure ποιότης : φέρεσθαι ἕκαστον τούτων ἅμα αἰσθήσει μεταξὺ τοῦ ποιοῦντός τε καὶ πάσχοντος καὶ τὸ μὲν πάσχον αἰσθητικὸν ἀλλ' οὐκ αἴσθησιν ἔτι γίγνεσθαι, τὸ δὲ ποιοῦν ποιόν τι ἀλλ' οὐ ποιότητα. « Translation de chacune d'elles et de la sensation correspondante dans l'intervalle situé entre l'agent et le patient ; le patient devient sentant, et non point sensation ; l'agent devient qualifié, et non point qualité. » Le dérivé est une inven-

[1] Extrait de la *Revue des études latines*, III, p. 214 et suiv.

tion peut-être de quelque sophiste, peut-être de Platon lui-même.

La traduction de la phrase de Platon que j'ai reproduite est excellente (c'est celle que l'abbé Diès a donnée à la collection Guillaume Budé) ; mais elle ne présente du texte qu'une image fortement déformée : pour Platon, pour l'Athénien qui lisait Platon au cinquième siècle avant Jésus-Christ, le mot ποιότης n'existait pas par lui-même ; ce n'était que le fait « d'être ποιος, d'être *quel* » ; pour nous, *qualité* est un terme fixé, qui ne se rattache à rien et que nous devons définir sans qu'aucun autre mot de la langue vienne l'appuyer. Le mot français qui est usé ne peut traduire le mot grec qui était neuf, et si avant que le traducteur soit entré dans son auteur, le français où il le fait passer ne lui a pas permis d'en rendre la valeur. On ne peut sentir Platon que dans son texte ; il faut une traduction pour aider à le comprendre, mais il est vain de le lire dans une traduction.

Si même *qualité* valait ποιότης, il n'en saurait avoir les résonances. Platon est l'un des créateurs de la pensée moderne, mais il est aussi un homme de son temps, encore engagé dans une pensée primitive. Pour lui, le mot n'a pas seulement son sens propre ; il évoque en même temps tous les mots dont la sonorité est voisine. Dans ce même passage du *Théétète*, l'auteur se sert de ποιότης en joignant le nom de notion à l'adjectif, dont l'abstrait est réellement un dérivé. Platon rapproche, on l'a vu, le verbe ποιεῖν, qui signifie « faire, et qui, pour nous, étymologistes exacts, n'a rien de commun avec ποιος et avec ποιότης. Le jeu avec les mots, qui se retrouve dans le même passage, 182 *b*, de manière claire, a eu d'ailleurs une large place chez Platon, comme on le voit par le *Cratyle;* ce jeu vient de la pensée primitive ; dans le peuple, il persiste aujourd'hui encore parmi nous, et les linguistes qui, comme Gilliéron, ont le sens intime de la langue populaire le sentent partout. Interprétant un texte grec, Eznik, un théologien arménien du cinquième siècle après Jésus-Christ, revient encore à traduire ποιότης, tantôt par un mot signifiant « qualité », tantôt par un mot signifiant « action » (cf. L. Mariès, *Le « Deo »*

d'Eznik, p. 116 et suiv. ; extrait de la *Revue des études arméniennes*, vol. IV et V).

De Platon à Aristote, il n'y a guère plus d'un siècle. Mais c'est un siècle où la philosophie, qui était, pour Platon, une recherche, est devenue la matière d'un enseignement systématique. Platon était un aristocrate qui pensait. Aristote a été un professeur. Pour Platon, ποιότης est une création avec laquelle il joue. Pour Aristote, c'est une catégorie qui figure à côté de la ποσότης, le fait d'être πόσος, la « quantité ». Le mot ποιότης est donc devenu un terme d'école. Mais il n'est encore qu'un terme d'école ; en grec, il n'a pas passé dans l'usage commun.

C'est Cicéron qui l'a lancé dans le monde.

Brillant élève de la philosophie et de la rhétorique grecques, et devenu par son talent un homme politique romain de rang élevé, Cicéron a constitué l'humanisme en faisant passer dans la langue des hautes classes de Rome ce que des gens du monde instruits peuvent acquérir de notions philosophiques.

Quand le régime césarien lui a fermé la bouche et a interrompu le rôle politique qu'il croyait jouer, il s'est mis à écrire des ouvrages de philosophie pour gens du monde. Il lui a fallu pour cela plier le vocabulaire latin à exprimer des idées pour lesquelles le latin n'offrait pas de ressources. La plupart du temps, il a usé de mots latins en les chargeant de sens nouveau. C'est ainsi qu'un mot comme *causa* n'est plus, chez Cicéron et ses contemporains, seulement le vieux latin *causa* ; c'est ce mot enrichi de tout le sens du grec αἰτία; depuis l'époque de Cicéron, et même un peu avant, la valeur de *causa* est toute transformée par l'interférence avec le grec αἰτία. Mais il y a des cas où, même ainsi sollicité, le latin ne fournissait pas l'équivalent des termes philosophiques grecs. Cicéron se hasarde alors, avec précaution, avec réserve, sans excès, à faire des mots latins que, en maître de la langue qu'il est, il forme bien : puisque le grec a fait ποιότης de ποιος, pourquoi ne pas faire *qualitas* de *qualis?* Cicéron l'a osé, et il a réussi. Il hasarde le mot dans le *De natura deorum*, dans les *Académiques;* il en donne d'abord l'équivalent grec ; il le répète ; il insiste à plusieurs re-

prises ; et, comme l'expression répondait à une idée, il l'impose. L'usage l'accepte.

Quoique dérivé de *qualis*, suivant un procédé bien latin, jamais *qualitas* n'a eu en latin la fraîcheur qu'a ποιότης chez Platon. Formé par un lettré, par simple transposition d'un terme grec d'école, *qualitas* a eu immédiatement la valeur d'un mot technique de la langue philosophique.

Voici le principal texte :

Acad., I, 24 : *Id iam corpus et quasi* qualitatem *quamdam nominabant; dabitis enim profecto, ut, in rebus inusitatis, quod Graeci ipsi faciunt, a quibus haec iam diu tractantur, utamur uerbis interdum inusitatis.*

Ayant ainsi proposé le mot, Cicéron se justifie, *id.*, 25 :

Nos uero, inquit Atticus, quin etiam Graecis licebit utare, cum uoles, si te Latina forte deficient. Bene sane facis; sed eniter ut Latine loquar, nisi in huiuscemodi uerbis, ut philosophiam aut rhetoricam aut physicam aut dialecticam appellem, quibus, ut aliis multis, consuetudo iam utitur pro Latinis. Qualitates *igitur appellant,* quas ποιότητας *Graeci uocant, quod ipsum apud Graecos non est ita uulgi uerbum, sed philosophorum, atque id in multis.*

Cicéron explique qu'il s'agit d'un terme technique et que chaque métier a le droit d'avoir les siens. Il ajoute donc, 26 :

Audebimus ergo nouis uerbis uti.

Il reprend ensuite, 27 :

... *Carentem omni illa* qualitate — *faciamus enim tractando usitatius hoc uerbum et tritius — materiam quamdam.*

Mais Cicéron n'a jamais employé comme un mot usuel ce *qualitas* qu'il ne lançait ainsi que comme substitut du mot aussi technique ποιότης. On le voit par le *De nat. deorum*, II, 94, où il écrit : *non* qualitate *aliqua quam* ποιότητα *Graeci uocant.*

Toutefois la différence qu'il y a entre la philosophie des écoles grecques et l'humanisme latin a produit ses effets. Le grec ποιότης n'est jamais sorti des écoles de philosophes, Cicéron le dit lui-même. Au contraire, *qualitas*, forgé, répété par un écrivain qui

était un homme du monde, a fait son chemin dans le grand public latin, et il est entré dans la langue des gens cultivés à Rome, sinon quand ils parlaient, du moins quand ils écrivaient. A la suite de Cicéron, tous les prosateurs ont employé *qualitas* librement.

Le parler courant ne possédait pourtant pas le mot *qualitas* ; les langues romanes n'en ont pas hérité, comme elles l'ont fait d'une autre création de Cicéron, *medietas* « μεσότης ». Mais le latin des livres était demeuré la langue de tous les gens qui avaient quelque culture. Tous les hommes cultivés de l'Europe y ont connu *qualitas*. En France, l'action de l'école a été grande dès le moyen âge. Et notre confrère M. A. Thomas nous apprend, dans le *Dictionnaire général*, que *qualité* et *quantité* ont été introduits dès le douzième siècle. On le voit, ces mots de la langue écrite ont été francisés sous la forme des abstraits, qui est celle de *bonté* et de *beauté*. Tout savants qu'ils étaient, *qualité* et *quantité* ont fait ainsi, dès l'abord, figure de mots français.

Mais, fait nouveau, et qui, du premier coup, distingue *qualité* de *qualitas*, ce *qualité* est en français isolé de tout autre mot. Le latin *qualis* était devenu *quel*, et, entre *quel* et *qualité*, un Français ne connaît plus le rapport que le Romain n'a pu cesser de sentir entre *qualis* et *qualitas*. Du coup, le mot *qualité* est libéré de tout lien avec sa formation ancienne. Il est autonome. Il n'est donc plus rien que le signe d'une idée qui, elle-même, n'est qu'un ensemble de caractères, et non pas une puissance active, comme l'idée platonicienne. Le langage a suivi ainsi le progrès de la pensée. Dans la langue des philosophes et des savants, la *qualité* s'oppose précisément à la *quantité*, et, grâce à ceci que ce terme ne fait partie d'aucune « famille de mots », *qualité* a la précision d'un terme technique dont le sens est enfermé tout entier dans les termes de sa définition et dont la netteté scientifique n'est troublée par les reflets d'aucun autre mot.

Mais, de tout temps, en France, le vocabulaire de la philosophie, de la science a été imité, reproduit dans le monde. Les mots du latin écrit sont entrés sans cesse dans l'usage courant.

Les juristes et les praticiens du droit, qui, comme les philosophes, ont un vocabulaire technique, ont pris le mot *qualité* ; quand on engage une procédure, quand on rédige un acte, le premier point est de déterminer à quel titre agissent les parties, de poser ce que l'on appelle les *qualités*. Et, comme les Français ont toujours eu le goût du droit strict, des pièces écrites, et que, d'ailleurs, les juristes ont toujours occupé une large place dans la grande bourgeoisie française qui a exercé sur la formation du français commun une action décisive, le mot est entré dans l'usage de tous, et l'on dit : *en qualité de*.

Par le fait qu'il vient de la langue écrite, *qualité* est un terme noble. Sauf indication contraire, une *qualité* est une « bonne qualité ». Sous l'ancien régime, les *gens de qualité* étaient de la classe noble.

Dès longtemps, les mots savants de cette sorte ont tendu à entrer dans la langue de tout le monde. Et maintenant il n'est personne qui ne parle des *qualités* d'un homme, de la bonne ou de la mauvaise *qualité* d'une marchandise. Ainsi le parler courant se nourrit de la langue des savants. Si la philosophie grecque n'avait fait ποιότης, si Cicéron n'avait, d'après le mot grec, fait *qualitas*, si les clercs n'avaient adapté *qualitas* pour en faire *qualité*, l'épicier ne parlerait pas couramment aujourd'hui de la *qualité* des produits qu'il vend.

Les autres langues romanes ont procédé comme le français, et l'italien, par exemple, a *qualità*, de même que le français a *qualité*. Mais c'est la forme française qui a servi de modèle dans l'Europe du Nord. C'est avec la forme *-té* du suffixe français que le moyen anglais a emprunté les abstraits latins en *-tas;* il en a fait *-tï*, et l'anglais moderne a *quality*. L'allemand a, de même, *qualität* depuis le seizième siècle.

Et le mot est si bien européen que les langues qui n'ont pas accepté le mot latin plus ou moins francisé l'ont traduit et transposé : *jakost* du tchèque, *jakość* du polonais ne sont toujours que le latin *qualitas*, sous une apparence tchèque ou polonaise. — Quant au russe *kačestvo*, c'est sur le grec qu'il repose ; car *kačĭstvo*

est attesté en russe dès le onzième siècle ; le mot fait partie de ces transpositions de termes techniques grecs qui sont nombreuses en slavon. Le mot est fait comme *qualitas*, mais il n'a dû sans doute sa fortune en russe qu'à l'importance prise par *qualitas* dans toutes les langues modernes de l'Europe.

Une création grecque, une adaptation latine, un intermédiaire français, ainsi se sont faits souvent les mots de la civilisation européenne.

LES INTERFÉRENCES ENTRE VOCABULAIRES[1]

Toutes les grandes langues de l'Europe moderne ont un mot pour dire « oui », et un Français d'aujourd'hui ne s'imagine pas qu'on puisse se passer d'un outil aussi familier. Pourtant ce mot n'est ancien nulle part ; le latin n'avait rien de pareil ; le grec non plus. Chaque langue romane a dû se donner un mot « oui » par ses propres moyens ; aussi le *oui* du français n'a-t-il rien de commun avec le *si* des Italiens ou des Espagnols. Un mot pour « oui » ne figurait pas parmi ceux que les peuples de langue indo-européenne ont reçus de la langue d'où sortent nos idiomes européens.

Qu'on ouvre les dialogues de Platon, on verra que la réponse à la question : « As-tu dit cela? » est volontiers : « Je l'ai dit » ou : « C'est moi (sous-entendu qui l'ai dit) », ou « Assurément ». Un Russe d'aujourd'hui à la question : « L'as-tu vu? » répond : « Je l'ai vu ». C'est que, en russe, la réponse *da* n'est pas idiomatique comme l'est, en français, celle par « oui ».

Que « oui » s'exprime par *si* comme en italien et en espagnol, par *ja* comme en allemand, par *yes* comme en anglais, par *tak* comme en polonais, par *da* comme en russe, il n'importe : partout dans les grandes langues de civilisation de l'Europe moderne, il y a, plus ou moins employé, un petit mot qui sert aux mêmes usages, qui rend à chaque instant les mêmes services.

[1] Lecture académique, octobre 1925.

Comment s'est réalisée cette unité nouvelle? on ne sait. Mais il semble bien que les langues se soient imitées l'une l'autre. Sans doute les sons diffèrent d'une langue à l'autre ; mais on peut imiter un procédé sans en reproduire le détail matériel. Un mot « oui » une fois créé dans une langue, ceux des voisins qui connaissaient cette langue auront à leur manière reproduit une pratique aussi commode.

Au premier abord, les vocabulaires de nos langues européennes sont très différents les uns des autres, et bien des gens s'efforcent de les rendre plus différents encore qu'ils ne le sont. Mais ces vocabulaires dont les sons ne concordent pas expriment un même fonds de civilisation. La science et la philosophie grecques, l'humanisme romain, le christianisme, la scolastique médiévale, la science expérimentale moderne nous ont fait une pensée commune et, que nous le voulions ou non, il nous faut des mots qui aient, sinon les mêmes sons, du moins les mêmes sens.

Quand on n'a pas emprunté des mots étrangers, il a donc fallu charger quelques mots de la langue d'exprimer les notions de la civilisation universelle, ou tirer de ces mots des formations qui permettent d'exprimer les notions communes.

A la fin du dix-huitième siècle, on a, en Allemagne, transposé *expression* en *ausdruck*, *impression* en *eindruck*. A première vue, l'allemand a ainsi l'air d'avoir des mots à lui, autres que ceux que les langues romanes et l'anglais doivent au latin. Pure apparence : sous un masque allemand *ausdruck* n'est toujours que *expression*, et *eindruck* que *impression*. Mais, au lieu des mots latins qui marquent l'unité d'une notion commune à tous les Européens, il y a des mots que l'étranger ne comprend que s'il les a appris. Et ces mots sont de moins bons signes que les mots latins devenus des mots européens. Car si *expression* et *impression* se rattachent à la notion concrète de « peser sur », ni un Français, ni un Anglais ne l'aperçoit : servant à exprimer des idées abstraites, ces vocables ont le mérite de n'évoquer rien de physique. Isolés de toute représentation concrète, ils sont des signes, nets comme les idées qu'ils traduisent. En allemand, *ausdruck* et *eindruck* évoquent *druck*,

qui est un nom verbal près du verbe *drücken*, et l'idée matérielle de « peser sur » risque de se mêler à la pureté de la notion abstraite.

Quand, au cours du dix-neuvième siècle, le mouvement démocratique a conduit à ériger en langues de civilisation des langues de paysans, on s'est piqué de tout dire au moyen des ressources propres de ces langues. On a poussé ce souci jusqu'à l'excès de bannir des termes universels du vocabulaire européen. En donnant au monde civilisé le modèle de la tragédie et de la comédie, les Grecs lui ont donné aussi le mot *théâtre*. Mais les arrangeurs du vocabulaire tchèque ont cru devoir le traduire par un mot nouvellement frappé, *divadlo* qui signifie littéralement « lieu où se donne un spectacle » ; double faute, car on perdait ainsi un terme européen et on le remplaçait par un mot qui, si on le prend au sens que sa formation suggère, exprime d'une manière vague une notion précise : tout endroit où se donne un spectacle n'est pas pour cela un théâtre. Pour sentir l'un des vices du procédé, il n'y a qu'à voir ce qui est arrivé au serbo-croate : on sait que Serbes et Croates se sont constitué une même langue littéraire, affirmant ainsi leur unité profonde et préparant, par la langue, l'union politique que la grande guerre a enfin réalisée. Mais, quand on a voulu rendre « théâtre » avec les ressources propres du serbo-croate, on l'a fait séparément en Serbie et en Croatie, et il y a eu deux mots différents : ce qui à Zagreb est *kazalište* est *pozorište* à Belgrade. Tant il est dangereux de vouloir, par des masques, dissimuler le mot *théâtre* que les Romains ont emprunté aux Grecs, et l'Europe aux Romains.

Les peuples modernes de l'Europe ont, par nationalisme, abusé du procédé qui consiste à transposer dans leurs langues des valeurs de mots étrangers. Mais le procédé est ancien, et c'est à des interférences de vocabulaires que nous devons beaucoup des termes que nous enployons à chaque instant.

Tout le monde parle de *cause*. Mais, si *causa* du latin a pris le sens de condition déterminante d'un fait, c'est parce qu'il a absorbé le sens du grec *aitia*. Au sens philosophique, les Romains n'ont employé *causa* que comme le substitut latin d'un mot grec.

Et *causa* est resté à Rome, en ce sens, un mot de la langue des gens instruits. Dans le parler du peuple, *causa* avait d'autres valeurs qui ont abouti à notre mot *chose*. Dans la langue écrite, chez les philosophes, chez les théologiens, *causa* a gardé l'équivalence une fois acquise avec le grec *aitia*. Le français a pris ce mot à la langue des clercs, en l'adaptant ; il a fait *cause*. Aujourd'hui *cause* est devenu un mot populaire, vulgaire même. L'enfant qui interroge en disant : *à cause?* ne parlerait pas ainsi si la philosophie grecque n'avait dégagé l'idée de la causalité, de l'*aitia*, et si les Romains n'avaient versé dans *causa* le contenu sémantique de *aitia*.

Ces interférences sont d'autant plus frappantes que les langues dont les vocabulaires ont interféré diffèrent davantage entre elles. La traduction de la Bible hébraïque en grec a donné à certains mots grecs des valeurs imprévues. Il sort de là dans le vocabulaire des peuples actuels de l'Europe des termes familiers à tous qui, de par leur origine, ne sont que des mots hébreux sous un déguisement grec.

L'habitude nous fait trouver naturel de désigner la divinité par le mot *Seigneur;* les Anglais disent de même *Lord*, les Allemands *Herr*, et ainsi partout. Or, nous n'en ferions rien s'il n'avait pesé à certain moment, sur l'un des noms hébraïques de Dieu, une interdiction. Dans la Bible, Dieu est nommé en partie d'un mot qui est noté par les quatre consonnes *y h w h;* c'est le tétragramme fameux qui, à un certain moment, est devenu ineffable. Partout où on le rencontrait, on le remplaçait par un substitut, par un mot signifiant « Seigneur ». Et ceci se marque dans la Bible hébraïque d'une manière expressive : quand, à une date tardive, les massorètes ont mis en dessous et en dessus des consonnes, seules écrites dans le texte primitif, des signes indiquant la prononciation des voyelles, les voyelles qu'ils ont juxtaposées aux consonnes du tétragramme *y h w h* ne sont pas celles qui auraient convenu à ce mot qu'ils ne prononçaient jamais, ce sont celles du mot signifiant « Seigneur » qu'ils prononçaient réellement ; c'est ce que des lecteurs naïfs ont lu *Jéhovah*, simple lecture d'ignorants. Les Juifs d'Égypte déjà hellénisés qui, vers le troisième

siècle avant Jésus-Christ, ont traduit la Bible en grec avaient dès lors l'usage de dire : « Seigneur » là où il y avait *y h w h*. Au lieu de transcrire en grec un nom qu'ils n'avaient pas le droit de faire entendre, ils ont traduit le mot qu'ils lisaient en fait, et c'est ainsi que le grec biblique *kyrios* « Seigneur » a désigné *Dieu*. En latin, *kyrios* a été traduit naturellement par *Dominus*, et le *Dominus* du latin par *Seigneur* en français, par *Herr* en allemand, par *Lord* en anglais, et ainsi de suite. Tous ces noms sont la traduction d'un nom grec qui, lui-même, traduit le substitut hébraïque du nom interdit *Jahweh*.

Des mots usuels de nos langues se sont produits par des traductions de cette sorte.

Ange est en français un mot courant ; il a l'air d'être grec, mais c'est à l'hébreu que nous le devons. Dans la Bible, Dieu, au lieu d'intervenir par lui-même, le fait souvent par des intermédiaires que le texte hébreu nomme *ml'k* « envoyé ». Dans leurs conceptions religieuses, les Grecs n'avaient rien de pareil. Les Juifs qui ont traduit la Bible en grec n'ont donc trouvé aucun mot grec pour rendre le terme hébraïque. Ils ont pris le nom usuel du « messager » qui était *angelos*. Dès lors, il y a eu en grec deux mots *angelos*, l'un universellement usuel, qui continuait de signifier « messager », l'autre propre à la langue biblique, qui résultait de l'interférence d'un mot hébreu avec un mot grec. Les auteurs du Nouveau Testament emploient l'un et l'autre mots : l'*angelos* proprement grec et l'*angelos* qui est le masque d'un mot hébreu.

Si fort qu'ils aient subi l'influence du vocabulaire grec, les Romains n'ont pas eu besoin d'emprunter le mot grec *angelos* au sens de « messager », ni l'occasion d'en subir l'influence : il y avait en latin ce qu'il faut pour désigner le « messager ». Mais, pour le sens biblique et chrétien, le latin n'avait pas plus de mot que le grec. Plutôt que de traduire *angelos* qui était clair seulement si l'on y ajoutait une détermination, si l'on disait *angelos kyriou*, « messager du Seigneur », le latin a pris l'*angelos* biblique. Et il a eu ainsi un nom biblique et chrétien *angelus* qui est emprunté au grec biblique. Mais, à la différence de *dominus*, ce mot *angelus* n'avait en latin

que son sens technique. Dès lors, on n'a pu le traduire ; les langues européennes n'ont eu d'autre ressource que de l'emprunter ; et toutes les langues de l'Europe ont hérité du latin *angelus*, en l'adaptant plus ou moins ; en français, l'emprunt est si ancien et l'adaptation si parfaite que *ange*, tout en gardant sa valeur technique, est devenu un terme du parler courant. Du mot hébreu, de valeur religieuse, à notre terme familier pour exprimer la tendresse il y a loin ; mais sans l'interférence qui s'est produite un jour entre un mot grec et un mot hébreu, la maman française ne dirait pas : *mon petit ange* à son enfant.

L'histoire du mot *diable* est plus compliquée ; mais elle commence de la même manière. Il y a dans la Bible un personnage mauvais qui dénonce les hommes à Dieu. C'est ce personnage qui sert à ouvrir le livre de Job. Job est un juste, mais tout lui réussit ; il est riche, il est heureux ; serait-il aussi pieux s'il lui arrivait malheur? Le *satan*, « le contradicteur », l'« opposant », demande à Dieu de tenter Job. Les Grecs ne connaissaient auprès des dieux aucun personnage malfaisant de cette sorte. Les traducteurs juifs de la Bible n'ont donc trouvé aucun terme propre pour traduire le *satan* hébreu. Mais, à côté du verbe *diaballein* qui ne signifie pas seulement « faire passer », mais aussi « tromper, calomnier, dire du mal de... », il y a normalement un nom d'agent *diabolos* « accusateur, calomniateur ». Ce mot n'était en grec ni usuel, ni bien important.

C'est un de ces vocables que les traducteurs sont heureux de trouver dans leur dictionnaire pour se tirer d'une situation embarrassante. Le *diabolos* joue dans l'Évangile un grand rôle ; il tente même Jésus. Pas plus à Rome qu'en Grèce, pareille figure n'était connue. Le latin biblique et chrétien n'a donc eu d'autre recours que d'emprunter le grec *diabolos*, et d'en faire *diabolus*. De là le mot est passé dans toutes les langues de l'Europe et l'anglais a *devil* aussi bien que le français a la forme francisée *diable*. Jusqu'ici, l'histoire est pareille à celle de l'ange.

Mais il s'est produit une seconde interférence. On sait quel rôle jouent dans l'Évangile les esprits malins qui pénètrent chez

l'homme et chez les animaux et leur font perdre la maîtrise d'eux-mêmes ; dans le texte grec de l'Évangile, ces personnages, dont le nombre n'est pas limité, sont nommés *daimones*, *daimonia*, d'un nom grec dévié de son sens propre. Ils n'ont rien à faire avec le *diabolos* qui est unique : le mot *satan* a fourni le nom propre *Satanas*. Mais on s'est mis à voir dans les *daimones* la contre-partie malfaisante des anges, et il s'est constitué ainsi comme des groupes ennemis d'anges et de démons. *Satanas*, le *diabolos* grec, le *diabolus* latin, est devenu une sorte de chef des démons. Et l'on a pu dès lors nommer *diables* les démons. Là où il n'y avait qu'un diable, il y en a des légions pour le peuple. Du coup, le mot français *diable* se trouvait avoir interféré avec *démon*, et à sa valeur ancienne, il avait joint un sens nouveau.

La valeur actuelle du mot *diable* en français et de ses équivalents dans les autres langues de l'Europe résulte donc d'une double interférence.

Ces exemples ont été pris parmi les plus clairs qu'on connaisse. Ils montrent combien complexe est l'histoire des mots. Plus on étudie l'histoire des langues, plus il apparaît que les actions qui se sont croisées sont multiples et diverses, plus aussi on voit que le parler du peuple se nourrit de la langue des savants. Quand on étudiait autrefois l'histoire des langues, on s'attachait à suivre la destinée des formes populaires d'une langue particulière au cours des siècles. Aujourd'hui, l'on sait que le vocabulaire de chaque langue est, souvent pour la plus large part, le produit d'influences étrangères et d'influences savantes. C'est la vieille civilisation égéenne qui nous a donné les noms de l'*huile*, du *vin*, de la *rose*, c'est la subtilité grecque qui nous a donné la *machine*, c'est la manière gauloise de faire la guerre qui nous a donné le *char*, c'est l'humanisme romain, héritier de la culture hellénique, qui nous a donné la *qualité*, pour ne rien dire de la masse des mots que nous devons aux écoles médiévales. Beaucoup plus qu'on ne le croit, beaucoup plus que ne le souhaitent des nationalismes myopes, les vocabulaires qui expriment notre civilisation européenne concordent entre eux. Nous devons à l'aristocratie indo-européenne

la base de notre organisation sociale qui a pris à Rome une forme nouvelle ; nous devons à la Grèce le système de notre pensée et à Rome l'adaptation de cette pensée à l'usage commun ; nous devons à la religion juive et à la religion chrétienne nos conceptions religieuses ; nous devons à la science expérimentale des derniers siècles nos idées sur le monde. Toute l'Europe actuelle, et l'Europe de langue germanique ou slave autant que l'Europe de langue romane, a hérité de ce fonds universel et, directement ou sous un déguisement, tous les vocabulaires intellectuels de l'Europe sont faits des mêmes éléments. Pour ce qui exprime la civilisation il y a dans nos langues, en dépit des amours-propres nationaux, beaucoup de bien commun, presque pas de bien particulier.

TABLE DES MATIÈRES

	PAGES
Avertissement	VI
L'état actuel des études de linguistique générale	1
Sur la méthode de la grammaire comparée	18
Note sur une difficulté générale de la grammaire comparée	36
Linguistique historique et linguistique générale	44
Convergence des développements linguistiques	61
Le problème de la parenté des langues	76
Les parentés de langues	102
Différenciation et unification dans les langues	110
L'évolution des formes grammaticales	130
Sur la disparition des formes simples du prétérit	149
Le renouvellement des conjonctions	159
Sur les caractères du verbe	175
Le genre grammatical et l'élimination de la flexion	199
La catégorie du genre et les conceptions indo-européennes	211
Comment les mots changent de sens	230
Le nom de l'homme	272
Quelques hypothèses sur des interdictions de vocabulaire dans les langues indo-européennes	281
A propos d'un récent dictionnaire étymologique du français	292
A propos des noms du vin et de l'huile	297
J. Gilliéron et l'influence de l'étude des parlers locaux sur le développement du romanisme	305
Sur le sens linguistique de l'unité latine	310
La religion indo-européenne	323
A propos de « qualitas »	335
Les Interférences entre vocabulaires	343

GLASGOW UNIVERSITY

L1-a27
General Singraf
A 150
ME1